

**DEVELOPMENT OF EMISSION GROWTH SURROGATES
AND ACTIVITY PROJECTIONS USED IN FORECASTING
POINT AND AREA SOURCE EMISSIONS,
FINAL REPORT**

**Prepared for the California Air Resources Board and the
California Environmental Protection Agency
Planning and Technical Support Division
Stationary Source Emission Inventory Branch
1001 I Street
Sacramento, CA 95814**

**Prepared by Andrew D. Bollman
E.H. Pechan & Associates, Inc.
3622 Lyckan Parkway, Suite 2002
Durham, NC 27707**

February 26, 2001

**Contract No. 98-719
Pechan Report No. 01.02.004/9425.000**

DISCLAIMER

The statements and conclusions in this report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

ACKNOWLEDGMENTS

This report was submitted in fulfillment of ARB Contract Number 98-719, Development of Historical and Forecast Growth Surrogate Data, by E.H. Pechan & Associates, Inc. under the sponsorship of the California Air Resources Board. Work was completed as of February 2001.

ABSTRACT

E.H. Pechan & Associates, Inc., and its subcontractors, Regional Economic Models, Inc. (REMI), and SDV-SCC Inc. have identified the most appropriate growth surrogates for point and non-mobile area sources for emission source categories defined by the California Air Resources Board (ARB). Historical data and projections of the surrogates were made under four growth scenarios for the years 1970 to 2030.

TABLE OF CONTENTS

DISCLAIMER	ii
ACKNOWLEDGMENTS	iii
ABSTRACT	iv
LIST OF FIGURES	viii
LIST OF TABLES	ix
SECTION I. INTRODUCTION	1
SECTION II. IDENTIFICATION OF GROWTH SURROGATES FOR POINT AND NON-MOBILE AREA SOURCE CATEGORIES	5
A. POINT SOURCES	5
1. Non-Fuel Combustion Sectors	5
2. Fuel Combustion Sectors	6
B. NON-MOBILE AREA SOURCES	8
SECTION III. DEVELOPMENT OF HISTORICAL AND FORECAST GROWTH SURROGATE DATA	11
A. REMI SOCIOECONOMIC DATA	11
1. REMI Model Overview	11
2. Forecast Data Developed from REMI's California Model	15
3. Data Anomalies	17
B. ENERGY ADJUSTED REMI SOCIOECONOMIC DATA	18
C. SOURCE CATEGORIES IDENTIFIED FOR DETAILED REVIEW	19
1. Selection of Source Categories	19
2. Emissions Activity Identification	19
3. Overview of Data Development Activities	20
4. Data Development Approaches for Each Source Category	24
Category 01. CES 46789 - Auto Refinishing	24
Category 02. CES 46748 - Industrial Coatings (Unspecified)	25
Category 03. CES 83030 - Organic Solvent Adhesives and Sealants	26
Category 04. CES 46441 - Oil and Gas Production, Other, Gaseous Fuel (Unspecified)	27
Category 05. CES 46458 - Oil and Gas Production, Crude Oil Production- Tanks	28
Category 06. CES 83089 - Consumer Products, Non-Aerosol Solvents	28
Category 07. CES 83550 - Agricultural Pesticides, Methyl Bromide	30
Category 08. CES 83568 - Agricultural Pesticides, Non-Methyl Bromide	32

TABLE OF CONTENTS (continued)

Category 09. CES 82123 - Residential Wood Combustion, Fireplaces	32
Category 10. CES 54569 - Residential Fuel Combustion-Space Heating, Natural Gas	34
Category 11. CES 54577 - Residential Fuel Combustion-Water Heating, Natural Gas	34
Category 12. CES 47332 - Agricultural Land Preparation	35
Category 13. CES 47340 - Farming Operations, Cattle Feedlot Dust	35
Category 14. CES 47357 - Building Construction Dust, Residential	37
Category 15. CES 47365 - Building Construction Dust, Commercial	38
Category 16. CES 47381 - Road Construction Dust	38
Category 17. CES 47399 - Unpaved Road Travel Dust-City and County Roads	39
Category 18. CES 47423 - Unpaved Road Travel Dust-Bureau of Land Management and Bureau of Indian Affairs Roads	40
Category 19. CES 47431 - Unpaved Road Travel Dust-Farm Roads	40
Category 20. CES 83337 - Fugitive Windblown Dust from Agricultural Lands (Non-Pasture)	41
Category 21. CES 83352 - Fugitive Windblown Dust from Unpaved Roads and Associated Areas	41
Category 22. SCC 10100601; SIC code 4931 - Electric Generation, Natural Gas, Boilers > 100 MMBtu except Tangential	42
Category 23. SCC 10100601; SIC code 4911 - Electric Generation, Natural Gas, Boilers > 100 MMBtu except Tangential	43
Category 24. CES 74682 - Cogeneration, Fuel Unspecified	43
Category 25. CES 82081 - I.C. Reciprocating Engines, Gaseous Fuel (Unspecified) (Oil and Gas Production)	43
Category 26. SCC 20200202; SIC code 1311 - I.C. Engines, Reciprocating, Natural Gas (Industrial)	44
Category 27. SCC 30600106; SIC code 2911 - Process Heaters, Process Gas-Fired (Petroleum Industry)	44
Category 28. CES 66787 - Internal Combustion (I.C.) Reciprocating Engines, Natural Gas (Manufacturing and Industrial)	44
Category 29. CES 47142 - Fuel Combustion, Other, Natural Gas (Manufacturing and Industrial)	45
Category 30. CES 83071 - Fuel Combustion, Other, Residual Oil (Manufacturing and Industrial)	45
Category 31. CES 83998 - Agricultural Irrigation, I.C. Engines, Diesel/Distillate Oil (Food and Agricultural Processing)	45
Category 32. CES 47233 - Fuel Combustion, Other, Liquid Fuel (Unspecified) (Food and Agricultural Processing)	45

TABLE OF CONTENTS (continued)

Category 33. CES 47167 - Fuel Combustion, Other, Natural Gas (Service and Commercial)	46
Category 34. CES 47159 - Fuel Combustion, Other, Distillate Oil (Service and Commercial)	46
Category 35. CES 57281 - Class II and III Landfills, Municipal Solid Waste ..	46
Category 36. CES 83659 - Degreasing, Cold Cleaning (Batch, Conveyor, Spray Gun), Petroleum Naphtha	47
Category 37. CES 58685 - Natural Gas Transmission Losses	48
Category 38. SCC 30500606 and SIC code 3241 - Cement Manufacturing (Dry Process)	48
Category 39. CES 85654 - Architectural Coatings, Oil-Based Industrial Maintenance Coating	49
Category 40. CES 85852 - Architectural Coatings, Water-Based Flat Coatings ..	50
Category 41. CES 46870 - Cutback Asphalt	50
Category 42. CES 66605 - Livestock Wastes	50
Category 43. CES 47241 - Agricultural Burning-Prunings	51
Category 44. CES 47258 - Agricultural Burning-Field Crops	51
Category 45. CES 47282 - Range Improvement	52
Category 46. CES 47274 - Forest Management	52
Category 47. CES 47266 - Weed Abatement	53
Category 48. CES 47290 - Non-agricultural Open Burning	53
Category 49. CES 60418 - Commercial Charbroiling	53
Category 50. SCC 20200202; SIC code 4922 - I.C. Engines, Industrial, Natural Gas, Reciprocating	54
D. ELECTRIC UTILITY-RELATED CATEGORIES	55
SECTION IV. SUMMARY AND CONCLUSIONS	57
A. SUMMARY OF RESULTS	57
B. STUDY LIMITATIONS AND RECOMMENDATIONS	60
SECTION V. REFERENCES	75
GLOSSARY OF TERMS, ABBREVIATIONS, AND SYMBOLS	80
TABLES REFERENCED IN REPORT	82
APPENDIX A. REMI ECONOMIC MODELING BACKGROUND MATERIALS	A-1
APPENDIX B. SAMPLE FRESNO GROWTH ACTIVITY PROFILE (GAP) DATA ...	B-1

LIST OF FIGURES

Figure IV-1. Residential Wood Combustion, Fireplaces	62
Figure IV-2. Agricultural Land Preparation	63
Figure IV-3. Unpaved Road Travel Dust - City and County Roads	64
Figure IV-4. Unpaved Road Travel Dust - Bureau of Land Management and Bureau of Indian Affairs Roads	65
Figure IV-5. Unpaved Road Travel Dust - Farm Roads	66
Figure IV-6. Fugitive Windblown Dust from Agricultural Lands (Non-Pasture)	67
Figure IV-7. Fugitive Windblown Dust from Unpaved Roads and Associated Areas	68
Figure IV-8. Class II and III Landfills, Municipal Solid Waste Landfills	69
Figure IV-9. Agricultural Burning - Field Crops	70
Figure IV-10. Residential Wood Combustion, Wood Stoves	71
Figure IV-11. Paved Entrained Road Dust -Local Streets and -Major Streets	72
Figure IV-12. Road Construction Dust	73
Figure IV-13. Comparison of Road Construction Dust Forecast Scenarios	74

LIST OF TABLES

II-1.	Unique List of Non-Fuel Combustion Point Source Growth Parameters	83
II-2.	Non-Fuel Combustion Point Source Growth Parameter Assignments	87
II-3.	Unique List of Fuel Projections Categories Incorporated Into REMI/AEO Growth Parameters	108
II-4.	Fuel Combustion Point Source Growth Parameter Assignments	110
II-5.	Non-Mobile Area Source Growth Parameter Assignments	164
III-1.	Details on Regression-Based Approaches	173
IV-1.	Number of Growth Parameters Included in Each Forecast Scenario	175
IV-2.	List of Growth Parameters Only Included in Best Estimate Forecast Scenario	176

SECTION I INTRODUCTION

Air pollution programs have always depended on predictive models for gaining a better understanding of what emission quantities will be in the future. The results of these models assist in the development of air quality plans; determine how and where air pollution can be reduced most efficiently; track progress toward meeting the requirements of air pollution control mandates; and are used to construct emission trends. California developed the Emission Forecasting System (EFS) in the 1990s to meet the multiple needs for this type of air pollution program planning information. A key component of California's EFS is the growth data.

The objective of this project was to identify the most appropriate growth surrogates for point and non-mobile area source emission categories within the State of California. A key component of this effort was to develop historical and projections data for these surrogates under alternative growth scenarios for the years 1970 to 2030.

Within California's EFS, growth indicators (or surrogates) can be assigned at various levels of detail. For example, indicators can be assigned at the 4-digit or 2-digit Standard Industrial Classification (SIC) code level. Area sources in California's Air Resources Board's (ARB) emissions inventory are identified by an emission inventory code (EIC), or a category of emission source (CES) code, while point sources are identified by a Source Classification Code (SCC) and a Standard Industrial Classification (SIC) code. Area source indicators are typically a mix of economic and non-economic indicators. An example of non-economic indicators for estimating agricultural PM emissions is the number of passes per acre per year by tilling equipment.

This study represented the first formal review of California's assignments of default growth indicators for each emission inventory source category. The information and data produced in this study results in significant improvements in the ability of the State to estimate area source emissions in historical and future years. Improvements in area source activity estimates and resulting emission estimates will make it more likely that California will develop effective control plans for ozone and PM_{2.5} precursors in future years.

The purpose of this report is to describe the efforts employed in identifying and developing emission growth surrogates for point and non-mobile area sources. E.H. Pechan and Associates, Inc.'s subcontractor, Regional Economic Models, Inc. (REMI) has developed economic models for each county in California. These models are based on historical data through the year 1997 and provide forecast data for four alternative growth scenarios (best estimate/most reasonable, optimistic/high growth, pessimistic/low growth, and cyclical) through the year 2035. These data provide socioeconomic indicators (e.g., employment and constant dollar output) for use as emission growth surrogates. For many point and non-mobile area source categories, Pechan assigned one of these REMI socioeconomic indicators to represent trends in emissions activity.

As part of this effort, Pechan has provided ARB with a list of growth surrogate assignments for each ARB point and non-mobile area source category. Pechan compiled this list in ARB's Parameter Assignment Data (PAD) file format. This list was developed based on a review of ARB's current growth surrogates, Pechan's experience in developing emission projections for various EPA regulatory efforts, and growth surrogate assignments from the United States (U.S.) Environmental Protection Agency's (EPA's) forthcoming Version 4.0 of the Economic Growth Analysis System (EGAS). (The EGAS employs REMI socioeconomic indicators as growth surrogates for most point and non-mobile area emission source categories.)

Under Tasks 2 and 3 of this contract, Pechan carried out a more detailed analysis of potential growth surrogates for 50 individual source categories. The source categories identified for detailed analysis were first determined based on source categories with the highest state-level 1996 criteria pollutant emissions. ARB staff then made some adjustments to the initial list based on other criteria that ARB wanted to consider (e.g., the "high profile" nature of certain source categories). The final list for detailed review contained 37 of the 50 highest-emitting categories and 13 categories specifically selected by ARB.

Under Task 2, Pechan conducted research into the availability of historical emissions activity data for these categories. This effort began with a review of ARB's base year emissions estimation methodologies. To the extent possible, Pechan applied ARB's methodologies in compiling a long-term historical data series for each of these 50 source categories. Under Task 3, Pechan researched the availability of forecasting data for these 50 source categories; for most emission activities, forecast data were not available. When forecast data were not available, Pechan conducted multiple regression analysis to identify socioeconomic variables from REMI's economic models that correlate with these emission activities. Based on the presence of a statistically significant relationship between the historical emissions activity data developed under Task 2, and one or more REMI socioeconomic variables, Pechan computed 1970-2030 data using the equation identified from the regression analysis. Pechan then reviewed these data to identify any data anomalies that necessitated adjustments to the regression-based approach. In a number of instances, the surrogate growth data indicated very large percentage changes in emissions activity over the forecast period. In such cases, Pechan reviewed the historical data used in the regression analysis to identify any explanations why the data series may not be representative of long-term trends. For example, the available 1990-1998 data on methyl bromide use indicates a dramatic drop in use over this period for structural applications. There were two developments in the 1990s that led to this uncharacteristic reduction – the Montreal Protocol and the California Department of Pesticide Regulation's strengthened pesticide use reporting requirements. Based on professional judgement, Pechan then adjusted the regression-based data to provide a more reasonable representation of long-term trends. Pechan compiled the final emission growth surrogate data into four sets of files in ARB's Growth Activity Profile (GAP) file format.

The remainder of this report is organized as follows. Section II identifies the growth parameter assignments for point and non-mobile source categories. This section also describes how Pechan compiled growth parameter assignment information in PAD file format. Section III discusses the REMI/Pechan methodologies for developing growth surrogate data for the 1970-2030 period for each of the growth parameters included in the PAD file. This discussion

includes the approach Pechan followed in preparing the four sets of growth surrogate data in GAP file format. A discussion of data sources used in developing these data is also provided in this section. Because ARB later identified particular interest in the growth surrogates assigned to electric utility-related sectors, this section includes a discussion of the growth surrogate data that will be used for these emission sources. Section IV presents a summary of the findings from this study, including sample graphs displaying major trends in the growth parameters assigned to the top twenty highest emitting source categories. This section also presents recommendations for potential future improvements. Section V identifies the references consulted for this study. All tables are located at the end of the report. Appendix A provides additional background materials on REMI's economic modeling methodologies. Appendix B presents sample growth surrogate data for Fresno county, which is located in the San Joaquin Valley Air Basin.

SECTION II

IDENTIFICATION OF GROWTH SURROGATES FOR POINT AND NON-MOBILE AREA SOURCE CATEGORIES

The purpose of this section is to describe the methods and rationale used in determining the most appropriate growth surrogates for ARB point and non-mobile area source categories. Pechan identified growth surrogate assignments based on a review of ARB's current surrogate assignments, Pechan's expertise related to preparing emission projections for various EPA regulatory efforts, and the set of growth surrogate assignments included in Version 4.0 of the EGAS, a growth surrogate projections tool Pechan has developed for EPA (Pechan, 2001). The growth surrogate assignments for point sources are described first. The point source discussion is organized into fuel combustion and non-fuel combustion sections. This is followed by a description of Pechan's approach for assigning growth surrogates to non-mobile area source categories.

A. POINT SOURCES

1. Non-Fuel Combustion Sectors

In developing the point source growth surrogate assignments, the initial question is whether to use the SCC, the SIC code, or a combination of the SCC and SIC code information in selecting an appropriate surrogate. Based on a review of the current ARB point source growth surrogate assignments, Pechan determined that the general ARB approach is to use SIC code information as the basis for the growth parameter assignment. For example, an inventory record classified in SCC 10100501--External Combustion Boilers, Electric Generation, Distillate Oil, Grades 1 and 2 Oil, currently uses output data for SIC code 24-Lumber and Wood Products as its growth parameter because the record is associated with SIC code 2499-Wood Products, Not Elsewhere Classified.

For this study, Pechan assigned growth surrogates using SIC code information unless data that are more specific to an SCC's emissions activity were available. In many cases, the SIC code of the point source records in ARB's inventory provides the best available information for projecting the emissions growth for that SIC code. In some instances, however, the growth in the projected SIC code will not capture many of the other factors that can impact the processes (SCCs) that directly relate to future emission levels. Fuel combustion-related SCCs, whose emissions are directly related to the amount of fuel (e.g. coal) burned, represent a clear example of this issue. The methods for identifying growth surrogates for point source fuel combustion sectors are described in the following section.

For this effort, Pechan employed SIC code output (total sales) data as the growth surrogate for all non-fuel combustion SCCs in ARB's point source inventory. Pechan used constant dollar

output data (in billions of 1992 dollars) rather than employment data because employment data will not capture productivity improvements or other changes that may affect the emissions activity per employee. In fact, the most recent guidance from the EPA's Projections Committee of its Emission Inventory Improvement Program states that the "employment level alone is not an effective growth indicator in most cases" (Pechan-Avanti, 1999). It is important to note, however, that absent a detailed analysis of trends in point source emissions, it is not possible to definitively state that output data correlates best with the emissions activity for each point source category.

Table II-1 displays the list of Pechan growth parameters used as growth surrogates for non-fuel combustion point sources in ARB's inventory. This table indicates the growth parameter name used in the PAD file, a description of the information it contains, and the SIC codes to which it is applied. Because REMI's models provide output data for 172 economic sectors, which are roughly equivalent to 3-digit SIC codes, Pechan directly matched each of the 1,082 valid 4-digit SIC codes to the economic sectors in REMI's models. In a few cases, REMI sectors are specified at other than the 3-digit SIC code level. For example, the Metal Mining sector is equivalent to SIC code 10 (REMI, 2000).

It is important to note that many SIC codes in ARB's 1996 emissions inventory do not represent a valid SIC code; often times these codes appear to be based on the outdated 1977 SIC code system. Whenever possible, Pechan used the first three digits of the invalid SIC code to match to the applicable REMI sector. For some point source records in ARB's inventory, SIC code information was invalid at the 3-digit SIC code level or available at less than a 3-digit SIC code level. For these cases, Pechan assigned a 2-digit SIC code-based REMI output indicator. For ARB inventory SIC codes that are also invalid at the 2-digit SIC code level (e.g., SIC code 2), Pechan used the SCC information in the inventory and the SCC-based REMI growth surrogate identified in the EGAS 4.0 model to assign a growth surrogate (EGAS employs REMI indicators as growth surrogates for most point source categories.) To ensure that these point sources are properly assigned growth parameters, ARB should consider reassigning valid SIC codes to the invalid/outdated SIC codes in its inventory.

The PAD file prepared by Pechan contains two types of point source growth surrogate assignments: fuel combustion and non-fuel combustion. The non-fuel combustion source categories are defined solely by 4-digit SIC code, and therefore are assigned an "8" to the "CATEGORY FLAG" field in the PAD file. Table II-2 displays the point source growth parameter assignments for non-fuel combustion source categories, including the information presented in the "PARAMETER COMMENT" field in the PAD file.

2. Fuel Combustion Sectors

For fuel combustion sectors, Pechan has assigned a separate set of point source growth surrogates. Although one would anticipate that the amount of fuel burned in a given industry is related to the amount of output that is produced by that industry, there are other factors that come into play. For example, fuel price changes can affect the amounts of different types of fuels burned. If, for example, the price of natural gas dropped significantly in relationship to other fuel types, one would anticipate that use of natural gas per unit of output would increase more than

the amount of other fuels. Another factor that lessens the correlation between output and fuel consumption is energy efficiency improvements, which can be required by government regulation, based on energy consumers' energy conservation decisions (e.g., insulation upgrades), and/or result from technology change. These price change and energy efficiency improvement factors are modeled in the Department of Energy (DOE), Energy Information Administration's (EIA) *Annual Energy Outlook* (AEO) fuel consumption projections (DOE, 1998). Unfortunately, DOE's set of projections do not provide State, county, or sub-county geographical detail, and therefore, are not directly used as growth surrogates in this effort.

Based on a similar approach used by EPA in preparing National emission projections, Pechan developed point source fuel combustion sector growth parameters from a combination of REMI socioeconomic data and fuel-specific energy consumption projections prepared for the AEO. In general this approach uses county-level SIC code-based output data from REMI's models adjusted for the national change in the intensity of energy use for each sector/fuel type as projected in the AEO. This approach acknowledges that emissions activities can grow at different rates within a given economic sector. Table II-3 presents a list of the AEO sector/fuel type categories used in developing REMI/AEO composite growth indicators in this study.

For this effort, Pechan reviewed the most recent list of point SCCs to which composite REMI/AEO parameters have been applied in projecting EPA's National Emission Trends inventory. Pechan revised this list based on the availability of more specific projections data for certain source categories. Specifically, Pechan assigned other growth parameters to certain source categories based on a more in-depth review of potential data sources that was conducted for certain significant source categories identified by ARB (e.g., electric utility fuel combustion sectors). The following briefly identifies the other data and methods used for these source categories (additional details on these data and methods are presented in Section III):

- *CATEGORY22* and *CATEGORY23* for point sources with SIC codes of 4911 and 4931 with natural gas combustion-related SCCs, Pechan used data on natural gas use by electric utilities. These data were available from the California Energy Commission (CEC) at the Air Basin/County level for 2000-2020 and at the state-level for 1976-1998. For 1970-1975, Pechan used state-level natural gas use by electric utilities as published by DOE's EIA. For 1999, Pechan assumed the average annual growth rate over the 1976-1998 period would hold for 1999-2000.
- *CATEGORY35* for Class II and III Municipal Solid Waste Landfills, Pechan used 1990-1999 county-level landfill tonnage data provided by the California Integrated Waste Management Board (IWMB).
- *NONGAS_UTIL* for point sources with SIC codes of 4911 and 4931 with fuel combustion SCCs, excluding natural gas, Pechan used data on non-natural gas fuel use by electric utilities. These data were available from the CEC at the Air Basin/County level for 2000-2020 and at the state-level for 1976-1999. For 1970-1975, Pechan used state-level non-natural gas use by electric utilities as published by the EIA. For 1999, a no growth assumption was applied based on the average annual 1976-1998 growth rate.

- *TOTAL_UTIL* for point sources with SIC codes of 4911 and 4931 with non-fuel combustion SCCs, Pechan used data on total fuel use by electric utilities. These data were available from the CEC at the Air Basin/county level for 2000-2020 and at the state-level for 1976-1999. For 1970-1975, Pechan used state-level fuel use by electric utilities as published by the EIA. For 1999, Pechan assumed the average annual growth rate over the 1976-1998 period would hold for 1999-2000.

As noted above, the PAD file compiled by Pechan contains two types of point source growth surrogate assignments - fuel combustion and non-fuel combustion. For fuel combustion categories, Pechan assigned growth surrogates based on SCC-SIC code combinations, and these categories, therefore, are assigned a "4" to the "CATEGORY FLAG" field in the PAD file. It is important to note that to reduce the potential size of the PAD file, Pechan developed the list of SCC and SIC code combinations based on the SIC codes included in ARB's 1996 emissions inventory that are assigned to fuel combustion SCCs. Table II-4 presents the point source fuel combustion sector growth parameter assignments. The REMI/AEO growth parameter names are a combination of the growth parameter name used for non-fuel combustion point sources and the fuel projection category codes presented in Table II-3 (e.g., SIC_206-I39 is a combination of the growth parameter SIC_206, which reflects output in the Sugar and Confectionery Products sector, and I39, which is the adjustment reflecting the projected change in steam coal use per dollar of output in the industrial sector).

B. NON-MOBILE AREA SOURCES

For non-mobile area sources, Pechan relied on a combination of growth surrogate assignment approaches. Most importantly, Pechan conducted an intensive search for historical and forecast emissions activity data for the non-mobile area source categories identified for detailed review by ARB (see Section III for a discussion of the categories included in this effort). For the ARB-identified categories with available historical emissions activity data, Pechan conducted multiple regression analyses. The purpose of the analyses was to identify the presence of a statistically significant relationship between the historical emissions activity data and potential growth surrogates. The analyses included potential growth surrogates from REMI's economic models and other data sources (e.g. on-road vehicle miles traveled data provided by ARB). Among the REMI model variables included in these analyses were output and employment data by economic sector, real disposable income, housing expenditures, and population. Based on the results of the analyses, Pechan identified an equation to be used along with the REMI data in developing growth surrogate data for certain non-mobile area sources. (For further details on these analyses, see Section III). To assign growth surrogates to the source categories not included in the detailed analysis, Pechan used REMI growth parameter assignment information provided in EGAS Version 4.0, the previous set of ARB growth parameter assignments, and professional judgment. Among the growth surrogates applied to these other categories were REMI output, employment, and population data, composite REMI/AEO data, and data supplied by ARB (e.g., on-road vehicle miles traveled data).

Based on previous ARB growth surrogate assumptions and time-series data obtained by Pechan, ARB directed Pechan to employ specific surrogate data and growth assumptions for

certain source categories. The following briefly identifies the other data and methods used for these source categories (additional details are presented in Section III):

- *CATEGORY04* for Oil and Gas Production, Gaseous Fuel (Unspecified), Pechan used California Division of Oil and Gas (DO&G) county-level natural gas production data to represent the 1970-1998 historical emissions activity trend in this source category; Pechan also used DO&G Northern and Southern California projections of natural gas production for every fifth year through 2022 in developing growth surrogate data through that year.
- *CATEGORY05* for Oil and Gas Production, Crude Oil Production-Tanks, Pechan used DO&G county-level oil production data to represent the 1970-1998 historical emissions activity trend in this source category.
- *CATEGORY07* for Pesticides/Fertilizers-Agricultural Pesticides-Methyl Bromide, Pechan used data on agricultural use of methyl bromide from the California Department of Pesticide Regulation (DPR) for 1990-1998, assumed growth through 2001 (based on the trend in Farm sector output), and then reduced the level of methyl bromide use to 30 percent of the 1998 DPR level for the year 2010 and all subsequent years.¹
- *CATEGORY08* for Pesticides/Fertilizers-Agricultural Pesticides (all except Methyl Bromide), Pechan used data on agricultural use of non-methyl bromide pesticides from the DPR for 1990-1998 to represent trends in the emissions activity for this category.
- *CATEGORY13* for Farming Operations, Cattle Feedlot Dust, Pechan employed ARB's Air Basin/County estimates of the number of cattle marketed in feedlots for 1987; Pechan also used State/District level cattle feedlot data to represent trends in use between 1970 and 1998.
- *CATEGORY37* for Natural Gas Transmission Losses, Pechan employed the data developed for *CATEGORY04* (see above).
- *POP-NOGROW* for Consumer Products-Aerosol Coatings categories, Pechan used population data as the growth indicator for all years except for 1991-2010 for the South Coast Air Basin and for 1991-1999 for all other Air Basins. For these years, Pechan employed a no growth assumption
- *STRUCTMETHYL* for Pesticides/Fertilizers-Structural Pesticides-Methyl Bromide, Pechan used data on structural use of methyl bromide from the DPR for 1990-1998 and

¹ ARB recommended that the 2010 values should represent one-third of the 1990 levels. Because the DPR data indicated that some areas' 1998 values are substantially below their 1990 values, applying this assumption for these areas results in substantial increases in agricultural methyl bromide use between 1998 and 2010. To alleviate this anomaly, Pechan calculated the State total use for 1990 and 1998 and developed the 30 percent reduction value to apply to the 1998 values based on the state-level increase in use that had occurred between those two years.

assumed straight-line reductions in use to a level of zero in year 2005 and all subsequent years.

- *STRUCTNONMET* for Pesticides/Fertilizers-Structural Pesticides (all except Methyl Bromide), Pechan used data on structural use of non-methyl bromide pesticides from the DPR for 1990-1998.

As requested by ARB, Pechan assigned growth surrogates to all non-mobile area sources based on EIC/SIC combinations. Therefore, all non-mobile area sources are assigned a "CATEGORY FLAG" of "9" in the PAD file. Table II-5 presents the non-mobile area source category growth parameter assignments, including the information provided in the "PARAMETER COMMENT" field in the PAD file. As noted in this field, it was necessary for Pechan to use state-level data in a few rare instances. In order to develop growth surrogate data for all counties under the regression-based approach, it was necessary for Pechan to use state-level regression-based activity data in cases where counties reported zero activity for a significant number of years or negative activity for at least one year.

SECTION III

DEVELOPMENT OF HISTORICAL AND FORECAST GROWTH SURROGATE DATA

This section describes how Pechan developed 1970-2030 data for each growth surrogate listed in Section II. Subsection A describes the development of REMI historical and forecast growth surrogate data. This is followed by subsection B, which discusses the development of composite growth surrogate data from REMI and AEO data sources. The historical and forecast data produced for each of the 50 source categories identified for detailed review is described in subsection C. This section concludes with subsection D, which presents an explanation of Pechan's methods for developing historical and forecast growth surrogate data for point sources identified with electric utility-related SIC codes.

A. REMI SOCIOECONOMIC DATA

For its model of the California county economies and aggregate State economy, REMI assembled data from public sources, calibrated the data to a published set of simultaneous equations, and generated forecasts based on nearly 30 years of historical data. The model is based on a full history from 1970 to 1997 for each county and includes a balanced history and calibrated forecast for each county for every year through 2030. Over 6,000 historical data points for each year are assembled in the model into an internally consistent time-series for each county. The REMI historical and forecast data for each county includes the following variables that are linked to emission source categories in Pechan's PAD file:

- 53 employment and output sectors broken out to 168 private sectors based on differential growth rates at the national level; and
- Population, gasoline and oil expenditures; and government employment.

In addition, several other REMI variables such as housing expenditures and real disposable income are used as input variables in the regression equation-based approach described in subsection C.

1. REMI Model Overview

REMI has developed the methodology used to build its socioeconomic models over the last 19 years. The REMI model is continuously improved by a team of researchers led by the creator of the modeling methodology and REMI's founder, Dr. George I. Treyz.

The REMI Economic and Demographic Forecasting and Simulation (EDFS) model used to prepare historical data and forecasts for the California ARB was customized to the economic and

demographic structures of California's counties. It includes State and county-specific data for industry-specific wage rates, production costs, employment, profitability, sales prices, consumer prices, housing prices, employment opportunity, population, State and local government spending, investment, income, personal consumption, and many other variables.

Underlying REMI U.S. Model

The REMI U.S. model provides the macroeconomic inputs to REMI models of California's 58 counties. The REMI U.S. model produces calibrated U.S. forecasts which act as drivers to the county models. Thus the impact of recessions, expansions, and other macroeconomic events are differentially and accurately reflected in the forecasts for each California county.

First documented in an article by Gang Shao and George Treyz in *Economics Systems Research*, Volume 5, No. 1, 1993, pp. 63-75, the REMI U.S. model incorporates full structural detail for the U.S. economy. It includes employment and output for 172 industries, projected technology changes, productivity projections, and endogenous responses to business fluctuations.

The basic structural parameters, including technological relationships, productivity projections, labor forces participation rates, final demand by sector, and other data are updated every two years based on the biennial detailed forecast issued by the Bureau of Labor Statistics. The model is designed so that alternative U.S. forecasts of final demands (for the 25 final-demand components of Gross National Product) are transmitted through the entire structure of the model to yield new projections for all of the variables in the model. REMI undertakes this updating annually using a forecast from the Research Seminar in Quantitative Economics (RSQE) from the University of Michigan, a leading provider of U.S. forecasts with one of the longest track records.

Model-based Forecasts

Several unique features of the California REMI integrated model make its long-term forecasts reliable bases for growth surrogates in California's large, complex economy:

1. The REMI model of California's counties is a *dynamic structural model* that captures the particular economic structure of each county and its interconnections to each of the other counties. These interconnections include trade flows by industry and commuter flows.
2. The model includes a highly developed *demographic component*, including a recently added Hispanic ethnic cohort. This dynamic detail is critical in California, given its ethnic diversity and differences in natality rates. The model's demographic module also explicitly includes migration from outside of the United States by ethnic group by county - a crucial growth driver in many counties in the State.
3. The REMI model independently defines a *labor force* for each of 160 population cohorts broken down by age, gender, and ethnic group. California's geographically and ethnically diverse population make this level of structural detail a key feature, given

differential participation rates across age, gender, and ethnic cohorts. These rates also respond to economic conditions endogenously.

4. The economic and demographic forecasts provided to the California ARB are linked in a theoretically consistent way: increased immigration to any county will lead, via lower wage rates due to expanded labor supply, to an increase in jobs, as the costs of doing business decline. On the other hand, an increase in demand from an external source will reduce unemployment and increase wages, which, in turn, will attract more people. These linkages are important to long-term forecasts and critical to forecasts based on alternative growth scenarios.
5. The REMI models' forecasts for California are based on key parameters estimated over large panel-data sets. Hence, the results escape the programs of regressions of short, regional data-series, which are subject to measurement error and large, random fluctuations due to non-predictable local events.

Model Structure

The structure of REMI's California economic model incorporates inter-industry transactions and endogenous final demand feedbacks. In addition, the model includes: substitution among factors of production in response to changes in relative factor costs; migration in response to changes in expected income; wage responses to changes in labor market conditions; and changes in the share of local and export markets in response to changes in regional profitability and production costs (REMI, 2000b). One strength of the REMI model lies in its use of theoretical structural restrictions instead of individual econometric estimates based on single time-series observations for each region.

The inclusion of price responsive product and factor demands and supplies give the REMI model much in common with Computable General Equilibrium (CGE) models. CGE models have been widely used in economic development, public finance and international trade, and have been more recently applied in regional settings. Static CGE models usually invoke market clearing in all product and factor markets. Dynamic CGE models typically assume perfect foresight inter-temporal clearing of markets, or temporary market clearing if expectations are imperfect. The REMI EDFs model differs, however, because product and factor markets do not clear continuously. The model replicates the time paths of responses between variables by combining *a priori* model structure with econometrically estimated parameters.

REMI models generate forecasts by solving a large number of simultaneous equations organized in five blocks as shown in Figure III-1, which describes the underlying structure of the model. Each block contains several components that are shown in rectangular boxes. The lines and arrows represent the interaction of key components both within and between blocks. Most interactions flow both ways indicating a highly simultaneous structure. Block 1, labeled output linkages, forms the core of the model. An input-output structure represents the inter-industry and

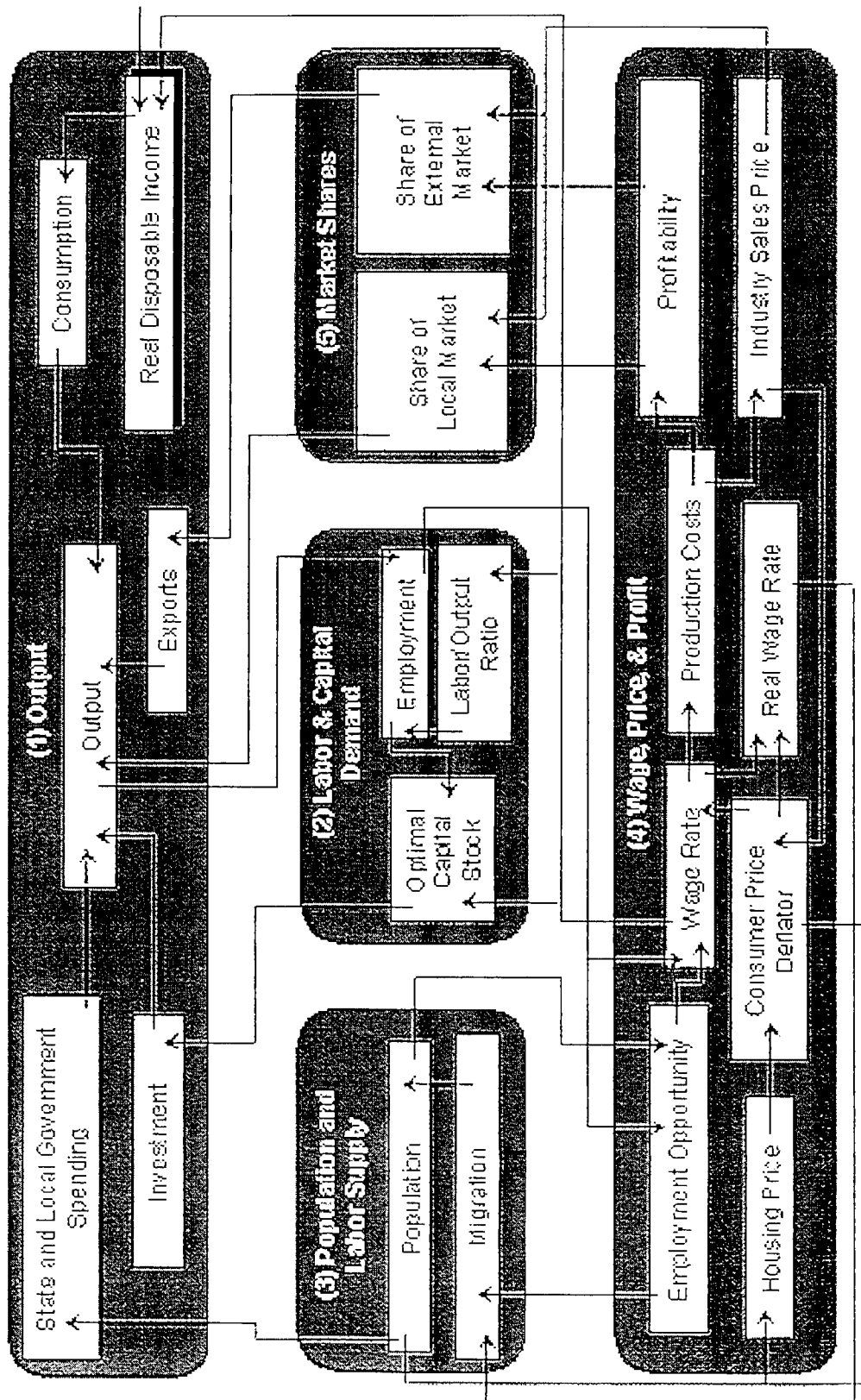


Figure III-1. Endogenous Linkages in the REMI Model

final demand linkages by industry. The interaction between block 1 and the rest of the model is extensive. Predicted outputs from block 1 drive labor demand in block 2. Labor demand interacts with labor supply from block 3 to determine wages. Combined with other factor costs, wages determine relative production costs and relative profitability in block 4 affecting the market shares in block 5. The market shares are the proportions of local demand in the region in block 1 and exogenous export demand that local production fulfills.

The endogenous final demands include consumption, investment, and State and local government demand. Real disposable income drives consumption demands. An accounting identity defines nominal disposable income as wage income from blocks 2 and 4, plus property income related to population and the cohort distribution of population calculated in block 3, plus transfer income related to population less employment and retirement population, minus taxes. Nominal disposable income deflated by the regional consumer price deflator from block 4 gives real disposable income. Optimal capital stock calculated in block 2 drives stock adjustment investment equations. Population in block 3 drives State and local government final demand. The endogenous final demands combined with exports drive the output block.

Appendix A presents detailed background materials on the historical data sources and methodologies used in developing REMI's California EDFS model.

2. Forecast Data Developed from REMI's California Model

REMI prepared forecasts for all the variables in the REMI model, according to four alternative growth scenarios. All forecasts were based on a REMI EDFS model of California's 58 county economies, calibrated with local historical data and integrated with each other to operate as a consistent model of the economy of the entire State (REMI, 2000c). Each forecast consisted of full, simultaneous solutions of the 58-area county model for each of the forecast years, reflecting the following four different growth scenarios:

1. A "standard" forecast (best estimate) scenario based on 20 years of model-based forecasting experience, extensive published reviews and the U.S. Bureau of Labor Statistics (BLS) and University of Michigan RSQE short and intermediate forecasts.
2. A "high growth" forecast (optimistic) scenario that used higher labor productivity assumptions, and included county-specific sales growth trends for the set of industries in the model that show higher growth than those incorporated in the REMI standard forecast.
3. A "low-growth" forecast (pessimistic) scenario that incorporated lower labor-productivity growth assumptions as well as demand trends not incorporated in the standard REMI forecast that would reduce the growth of the counties in question.
4. A "cyclical" forecast that is based on business cycle assumptions applied to the underlying U.S. model, which acts as the macroeconomic driver to the models of the California counties. The cyclical forecast reflects fluctuations in final demand due to recessions and recoveries, as well as changes in labor productivity that endogenously

respond to changes in the rate of output growth by industry in the sample period. Thus, the California multi-county forecast reflects the cyclical events that are predicted in the REMI U.S. cyclical forecast.

REMI used the following specific procedures to develop its three alternative economic forecasts.

High and Low Economic Growth

For these two alternative forecasts, REMI modeled significant historical employment trends that are not explained when REMI back-casts employment using its model structure. Under the high economic growth forecast, REMI included employment trends that were higher than explained by REMI's modeling approach; for the low economic forecast, REMI included employment trends that were lower than expected given REMI's model structure. REMI included employment trends by county and industry when judged significant and persistent over the last ten years.

In addition, the high growth forecast included an increase in productivity growth of 0.6 percent per year above the projections in the REMI standard forecast; the low growth forecast included a decrease in productivity growth of 0.6 percent per year versus the standard forecast. It should be noted that the decrease in productivity growth was not applied to sectors that had little or no productivity growth under the standard forecast. Because the REMI model treats the farm and federal government sectors differently from other economic sectors (e.g., farm employment and value added are based on the local share of U.S. farm employment and value added), REMI established the high and low growth forecasts for these sectors in a different manner. The extent of the adjustment for each sector and forecast was developed based on the nature of the sector, its movements over the last 30 years and judgement about the likely deviation from the standard U.S. forecast.

Cyclical Economic Growth

For this alternative forecast, REMI first measured the variance around the national trend in each of the following final demand variables over approximately the last 30 years:

- Fixed residential investment;
- Fixed non-residential investment;
- Producer's durable equipment;
- Farm;
- Government (Federal, civilian; Federal, military; State and local); and
- Exports.

REMI then superimposed the identified cyclicalities in each variable to the REMI standard U.S. forecast over the 2001 to 2035 period. This approach reflects all of the major cyclical factors that are not captured endogenously by the U.S. model (e.g., changes in labor productivity that endogenously respond to changes in the rate of output growth by industry). Because the U.S. model acts as the macroeconomic driver for the California county economic models, the changes in the U.S. forecast result in cyclical changes to the California county economies.

3. Data Anomalies

Historical Data

In some instances historical data in REMI's models indicate anomalous year-to-year trends in employment or output. For example, output in the child day care services sector increases from \$376 million to \$915 million between 1971 and 1972. REMI reviewed the historical data and determined that there are two explanations for these anomalous values, each of which result from data limitations: (1) the need to disaggregate available Bureau of Economic Analysis (BEA) one-digit SIC code county-level BEA employment data to the two-digit level; and (2) data suppressions in the BLS employment and output data series. For the first case, two-digit SIC code data are constructed from the BEA one-digit SIC code county series using two-digit SIC code county data from the BLS ES-202 series and the Department of Commerce's County Business Patterns. Sharp rises or declines in the REMI county-level employment and output data are attributable to a similar fluctuation in the ES-202 data or the use of County Business Patterns values where the ES-202 numbers were suppressed. For the second case, significant changes can occur from year-to-year when there are years with more data suppressions than other years.

It is important to note that the economic sectors where the data anomalies occur are not matched to major emitting source categories. For example, the largest year-to-year changes occur with the following sectors/variables:

- output in job training and related services;
- output in museums, botanicals, and zoological gardens;
- output and employment in tobacco manufacturing/products;
- output in forestry, fishing, hunting, and trapping;
- output in residential care; and
- output in miscellaneous fabricated textile products.

In addition, the majority of data anomalies occur in the earliest years of the study period. In particular, many of the anomalies result from an unusually high number of data suppressions prior to 1972.

Forecast Data

It should be noted that the high- and low-growth forecasts sometimes result in employment and output values that appear counterintuitive when compared to the best estimate (standard) forecast values. For example, under the low-growth forecast for 2030, employment in the railroad equipment sector is approximately 30 percent greater than employment under the standard forecast. The explanation for this occurrence is that labor costs in the counties where railroad activities take place are projected to decline under the low-growth scenario. This cost reduction is the result of a decline in employment for other industries that use the same labor force skills as the railroad equipment sector. This drop in the local cost of production results in increased local competitiveness in this sector relative to the rest of the nation. Because railroad equipment demand is forecast for only a small decline/slight increase under the low-growth scenario and because U.S. labor productivity is reduced by 25 percent in the low growth scenario,

the result is a higher level of projected railroad equipment employment versus the level projected under the standard forecast.

B. ENERGY ADJUSTED REMI SOCIOECONOMIC DATA

Many ARB source categories are associated with fuel combustion emissions activities. Because of the possibility that emissions activities will grow at different rates within a given economic sector, and the availability of fuel specific projections from the AEO, Pechan developed composite projections data for fuel combustion related point sources. These composite projections are developed from a combination of REMI socioeconomic data and AEO energy consumption data (REMI, 2000c and DOE, 1998). Pechan compiled the available AEO data for 1996-2020 and developed adjustment factors relating to changes in energy intensity. Specifically, Pechan calculated the following national energy intensity factors for 1996-2020:

- Residential fuel combustion - projected delivered energy by fuel type divided by the projected number of households;
- Commercial/institutional fuel combustion - projected delivered energy by fuel type divided by total Services sector output in millions of 1987 dollars; and
- Industrial fuel combustion - projected delivered energy by fuel type for both specific industries (e.g., refining industry) and for total industrial fuel use divided by projected industrial sector output in millions of 1987 dollars (specific industry or total industrial output).

Next, Pechan calculated the ratios of national 1996 energy intensity to the national energy intensity for each sector/fuel type through 2020. For example, AEO projects natural gas consumption in the commercial sector to rise from 3.392 quadrillion Btu in 1996 to 3.997 quadrillion Btu in 2020. Over this same time-frame, AEO projects constant dollar output in the Services sector to increase from approximately 6.2 to 10.7 billion dollars (in 1987 dollar terms) between 1996 and 2020. To reflect the projected change in natural gas consumed per dollar of Services sector output, natural gas energy intensity factors were calculated for 1996 and each projection year. For example, 0.375 quadrillion Btu/per dollar of natural gas is projected to be consumed in 2020 versus 0.545 quadrillion Btu/per dollar in 1996. For all commercial sector natural gas source categories in 2020, the REMI commercial sector growth factors are multiplied by 0.69, which represents the ratio of the 2020 energy intensity factor for commercial sector natural gas to the 1996 energy intensity factor for commercial natural gas. Similar ratios were calculated and applied for other fuels energy sectors and projection years.

Because the AEO forecast data ends in 2020, Pechan projected energy intensities through 2030 based on historical trends from the AEO projections series.

C. SOURCE CATEGORIES IDENTIFIED FOR DETAILED REVIEW

1. Selection of Source Categories

Pechan conducted a detailed review of growth surrogate assignments for 50 source categories identified by ARB. These 50 categories were identified from a mix of the highest-emitting categories in ARB's 1996 emissions inventory (37 categories) and ARB preferences (13 categories).

Pechan identified the top 50 emitting source categories from the ARB 1996 emission inventory by totaling the individual pollutant emissions for each source category. Emissions for the following pollutants were totaled for this effort: reactive organic gases (ROG), oxides of nitrogen, carbon monoxide, oxides of sulfur, and particulate matter with an aerodynamic diameter of 10 micrometers or less (PM₁₀). The list of top emitting source categories was developed at the SCC/SIC code-level for point sources and the Category of Emission Source (CES)- level for area sources. The resulting top 50 source category list accounts for approximately 83 percent of total point and non-mobile area source emissions in ARB's 1996 inventory. ARB then reviewed this list and identified changes eliminating 13 of the top 50-emitting source categories and replacing these with different categories selected by ARB. The purpose of this subsection is to provide details on the methods Pechan employed in developing historical and forecast emissions activity data for each of the final 50 categories. Because of recent developments in the electric utility industry, ARB subsequently requested that Pechan focus additional efforts on developing growth surrogate data for electric utility SIC codes (4911 and 4931). Subsection D describes the data developed for SIC codes 4911 and 4931.

2. Emissions Activity Identification

To begin the effort of developing historical emissions growth surrogate data for the ARB identified categories, Pechan reviewed three main sources of emissions activity information:

- "Emission Inventory 1996," State of California, Air Resources Board, Technical Support Division, Emission Inventory Branch. October 1998 (ARB, 1998a);
- "Index of Methodologies by Major Category," posted on ARB's web site at <http://www.arb.ca.gov/emisinv/areasrc/index0.htm> (ARB, 2000); and
- "Emission Inventory Procedural Manual, Volume III, Methods of Assessing Area Source Emissions," State of California, California Environmental Protection Agency, Air Resources Board, October 1997 (includes all revisions through November 1998) (ARB, 1997).

Pechan reviewed these sources for information on the methodologies and data sources employed by ARB in developing base year emission estimates for each of the emission source categories identified for detailed review. Pechan also requested and received additional background documentation of these methods and data sources from ARB. For some categories, ARB documentation was not available because these categories' emissions are not estimated by ARB,

but instead, are estimated by each Air District. Examples of such source categories include CES code 66605–Livestock Wastes and CES code 60418–Commercial Charbroiling. For source categories that are the responsibility of Air Districts, Pechan reviewed EPA information sources and any readily available District documentation to identify activities likely to be related to these categories. For the Agricultural Burning–Field Crops category (CES code 47258), for example, Pechan identified a list of crop types with emission factors in EPA’s *AP-42* document and used this list in compiling crop acreage data for use as emissions activity data for this source category (EPA, 1999). Pechan also reviewed the latest emissions inventory preparation guidance for sources of information for historical and forecast emissions activity data, and the factors affecting the emissions activity. This guidance is available from EPA’s Emission Inventory Improvement Program (EIIP) web site. The EIIP guidance was reviewed for the following area source categories:

- Residential Wood Combustion (Radian, 1997);
- Architectural Surface Coating (Radian, 1995);
- Consumer and Commercial Solvent Use (ERG, 1996);
- Solvent Cleaning (ERG, 1997a);
- Industrial Surface Coating (TRC, 1997a);
- Pesticides-Agricultural and Nonagricultural (TRC, 1997b);
- Asphalt Paving (ERG, 1998);
- Open Burning (ERG, 1999);
- Municipal Landfills (ERG, 1997b); and
- Autobody Refinishing (ERG, 2000).

In a few rare cases, no supporting emissions activity documentation was identified. Pechan identified emissions activities for these categories based on professional judgement.

3. Overview of Data Development Activities

After identifying the emissions activity data for each source category, Pechan researched the availability of long-term historical data for these categories beginning with 1970 (the first year of the time frame for this study). After obtaining and compiling historical time-series data for these categories, Pechan researched the availability of forecasts for these emissions activities. In the vast majority of cases, Pechan was unable to identify such forecasts. For these cases, Pechan conducted multiple regression analyses to identify surrogate indicators that correlate with the historical emissions activity data. Pechan selected potential explanatory variables for these regression analyses based on the EIIP emissions inventory preparation guidance, the availability of data from REMI’s economic models, and professional judgement. Pechan selected growth surrogates for use in projecting emissions based on the identification of statistically significant equations from the regression analysis. Pechan conducted the regression analyses using state-level historical emissions activity and REMI data.

Table III-1 displays information about the regression equations employed, including the variables identified as statistically correlated with historical emissions activity data and two important statistics (adjusted r^2 and t-statistic) that indicate the strength of the correlation

between the surrogate indicator(s) and the emissions activity. For source categories with more than one statistically correlated variable, the most strongly correlated variable is identified in boldface type. This table also indicates the base year of the data employed in each source category's regression analysis.

Pechan then employed the REMI best estimate forecast State data into the regression equations to develop state-level projections data throughout the study period. Pechan then adjusted the resulting state-level values for each county's growth relative to State growth. For the county adjustments, Pechan computed factors reflecting the county change relative to the state-level change in the most strongly correlated REMI variable identified from the multiple regression analysis. For those counties and variables that REMI shows with zero activity for certain years, Pechan directly used the state-level regression output to represent the county-level trend. Pechan developed the county-level adjustments using the following two algorithms.

Algorithms for State-level Growth Surrogates with Projected Increases Between 1970 and 2030:

$$C_{ify} = S_{ify} * [(c_{ify} - c_{iby}) / (s_{ify} - s_{iby})]$$

$$C_{ity} = S_{iby} + [(C_{ify} - S_{iby}) * \frac{(c_{ity} - c_{iby})}{(c_{ify} - c_{iby})}]$$

Algorithms for State-level Growth Surrogates with Projected Declines Between 1970 and 2030:

$$C_{ify} = S_{ify} * [(s_{iby} - s_{ify}) / (c_{iby} - c_{ify})]$$

$$C_{ity} = S_{iby} - [(S_{iby} - C_{ify}) * \frac{(c_{ity} - c_{iby})}{(c_{ify} - c_{iby})}]$$

where:

- C = county growth surrogate value;
- S = state regression output;
- c = county growth index calculated from data for most strongly correlated variable;
- s = state growth index calculated from data for most strongly correlated variable;
- i = source category;
- ty = target year of calculation;
- by = 1970; and
- fy = 2030.

Sample calculations for the Automobile Refinishing growth parameter in Fresno county are described to illustrate the procedure. As indicated in Table III-1, the multiple regression analysis identified employment in the Automobile Parking, Repair, and Services sector (SIC_752-4emp) as the sole statistically significant correlate with the emissions activity data for this source category. As indicated by the identified equation, the analysis indicated a 19.4 percent increase in emissions activity with every 100 percent increase in employment in this sector. The first step in the procedure is to compute indices representing the change in state-level employment in the

Automobile Parking, Repair, and Services sector for the base year of the data used in the regression analysis for this category (1987). State level SIC_752-4 emp data and calculated indices are shown below.

	1970	1987	2030
REMI data	61.629	147.106	375.634
Index (1987=1)	0.419	1.000	2.553

The next step is to compute the output from the identified regression equation by inputting the state-level indices computed above.

	1970	1987	2030
SIC_752-4emp index (1987=1)	0.419	1.000	2.553
Equation	$y = .801 + .194 * .419$	$y = .801 + .194 * 1$	$y = .801 + .194 * 2.553$
Equation output	0.882	0.995	1.296

The next step is to compile Fresno county data representing employment in the Automobile Parking, Repair, and Services sector analogous to the state-level data presented earlier.

	1970	1987	2030
Fresno county SIC_752-4emp	1.508	2.622	6.769
SIC_752-4emp index (1987=1)	0.575	1.000	2.582

The next step estimates the 2030 Fresno county value using the first equation listed on the previous page.

$$\begin{aligned}
 \text{Fresno final data for 2030} &= 1.296 * [(2.582-.575)/(2.553-.419)] \\
 &= 1.296 * (2.007/2.134) \\
 &= 1.219
 \end{aligned}$$

The next step estimates the Fresno county values for 1971-2029 based on the second equation listed on the previous page. This step uses the 1970 State-level regression output, the 2030 Fresno estimate displayed above, and the ratio of Fresno county SIC_752-4emp growth between 1970 and the target year and Fresno county SIC_752-4emp growth between 1970 and 2030.

The purpose of the final step is to adjust the state-level regression output to reflect the relative growth in employment in the Automobile Parking, Repair, and Services sector in Fresno county versus the growth in this sector at the State level. Below shows a sample calculation for year 1987.

	1970	1987	2030
State equation output (from above)	0.882	0.995	1.296
County adjustment calculation for 1987		$= .882 + (1.219 - .882) * \\ [(1 - .575) / (2.582 - .575)]$	
Final county data	0.882	0.953	1.219

Pechan then reviewed the output from implementing the appropriate algorithms to identify the reasonableness of the resulting growth rates. Pechan revised some of the initial county-level growth surrogate estimates to temper the more extreme county values. To systematically identify the estimates to revise, Pechan calculated the 1970 to 2030 county-level growth rates for each growth parameter and the standard deviation of these county growth rates. Pechan modified the initial growth surrogate estimates for all counties whose 1970-2030 growth rate was more than one standard deviation above or below the average 1970-2030 growth rate for a given parameter. In nearly all instances, Pechan set the revised county 1970-2030 growth rate equal to one standard deviation away from the average category growth rate. For example, counties whose growth rates were more than one standard deviation below the average 1970-2030 growth rate for a given parameter, had their 2030 estimate revised to equal the growth representing one standard deviation below the average county growth rate. The 1971-2029 growth surrogate data were then calculated by apportioning the 1970-2030 rate-of-change using the initial growth surrogate estimates. Specifically, Pechan multiplied the revised 1970-2030 change by the ratio of the change between the initial 1970 surrogate value and the initial target year value to the change between the initial 1970 and 2030 values. This algorithm is similar to the one described earlier in relation to computing the initial county-level growth surrogate estimates. In the small number of cases where this approach yielded a negative estimate for one or more years, Pechan replaced the entire initial county level growth surrogate series with the state-level regression series.

For some source categories, Pechan determined that the regression-based approach provided large growth rates that appear unsustainable in the long-run. For all but three of these source categories, Pechan employed the regression-based approach only for the years for which the regression analysis was conducted. For the remaining years over the study period, Pechan applied the county-level growth rate for the variable identified as the most strongly correlated with the historical emissions activity data.

For three source categories – Industrial Coatings (Unspecified); Architectural Coatings, Oil-Based Industrial Maintenance Coatings; and Structural Methyl Bromide – Pechan did not use the regression-based approach for any years, but instead used the county-level trends in the variable identified as most strongly correlated with the emissions activity data. Data limitations explain

why the regression equations developed for these three source categories did not provide reasonable growth surrogate values. For the Industrial Coatings (Unspecified) category, only National time-series data are available, and the unclear definition of this category makes the available data questionable for determining long-term trends in California emissions activity for this category. Similarly, no California-specific time-series emissions activity data were identified for the Architectural Coatings, Oil-Based Industrial Maintenance Coatings category. In addition, National production data specific to industrial maintenance paints were only available for 1982-1984. For the Structural Methyl Bromide source category, Pechan utilized 1990-1998 DPR data on the amount of methyl bromide used in structural applications. The DPR data indicate that the amount of methyl bromide used in these applications declined from 5.2 million pounds in 1990 to 363 thousand pounds in 1998. This dramatic reduction in methyl bromide use resulted from DPR's additional pesticide reporting requirements and the impact of the Montreal Protocol. Estimating the pre-1990 trend in structural methyl bromide use based on the identified equation yields unreasonably large values. The regression-based approach to back-cast the 1990 DPR data is not valid because of the policy changes that uncharacteristically influenced the trend in structural methyl bromide use over the regression analysis period.

The following presents details on the data sources and approaches Pechan employed in developing historical and forecast data for each of the 50 ARB-identified source categories, and provides additional information on the data limitations associated with these categories.

4. Data Development Approaches for Each Source Category

Category 01. CES 46789 - Auto Refinishing

ARB's "Emission Inventory Procedural Manual, Volume III, Methods of Assessing Area Source Emissions," hereafter referred to as ARB's Methods Manual, describes the use of National production data for "auto and machinery refinish paints" from the Bureau of the Census and vehicle registration data to estimate state-level base year emissions activity data for this source category (county emissions activity estimates were based on county to State population ratios). To develop historical time-series data for this category, Pechan first compiled National "auto and machinery refinish paints" production data from the Bureau of the Census (Census, 1997a and Census, 1997b). Because of breaks in the available data series, Pechan compiled these data for 1972, 1977, and 1982-1997. Pechan then employed these data in a multiple regression analyses. To test potential indicators that correlate with production of automobile and machinery refinish paints, Pechan regressed these data against the National historical values for the following variables: employment and output in the Automobile Parking, Repair, and Services sector; vehicle miles traveled (VMT); and population. With the exception of the VMT data, Pechan obtained the values for these indicators from REMI's "best estimate/standard" U.S. economic forecasting model. The National VMT data were compiled from EPA's MOBILE 4.1 Highway Fuel Combustion Model. It should be noted that VMT, population, and employment in the Automobile Repair Services sector were identified as indicators of automotive refinishing activity in the EIP guidance for Autobody Refinishing, the SRI International report entitled, "U.S. Paint Industry Data Base," and/or ARB's Methods Manual (ERG, 2000; SRI, 1990; ARB, 1997). None of the indicators explained a large proportion of the trend in the underlying Bureau of the Census data series. The best statistical fit identified was for Automobile Parking and

Repair Services employment (adjusted $r^2=0.353$, t-statistic = 3.203). Pechan believes that the underlying Bureau of the Census data may not provide the best indication of automotive refinishing activity because these data include "other transportation and machinery refinish paints."

Pechan, therefore, developed a multiple regression analysis based on available historical and forecast data for National motor vehicle refinishing coatings demand data reported in the Freedonia Group, Inc.'s report "Paints and Coatings to 2002" (Freedonia, 1998). This regression analysis indicated a very strong correlation, and the best statistical fit, with Automobile Parking and Repair Services employment ($r^2 = 0.949$; r^2 adjusted = 0.932; t-statistic of 7.464). Pechan prepared the growth surrogate data for this source category based on a combination of REMI state and county-level employment data for the Automobile Parking, Repair, and Services sector and the regression-based approach described in Section III.C.3.

Category 02. CES 46748 - Industrial Coatings (Unspecified)

The following Bureau of the Census publications provide the main emissions activity data for this source category: *Annual Survey of Manufacturers*, "Paint and Allied Products," and *Current Industrial Reports*, "Paint, Varnish, and Lacquer" (Census, 1997a and Census, 1997b). ARB estimated 1982 emissions for this category by summing the following Census categories:

- 28512 23 - Truck and bus finishes;
- 28512 25 - Other transportation equipment finishes, including aircraft, rockets, and missiles;
- 28512 33 - Railroad finishes;
- 28512 35 - Appliance, heating equipment, and air conditioner finishes;
- 28512 39 - Wood and composition board flat stock varnishes and other finishes;
- 28512 53 - Insulating varnishes, electrical types;
- 28512 58 - Other industrial product finishes, excluding semimanufactured products such as pigment dispersions and ink vehicles;
- 28512 65 - Industrial lacquers, fabricated metals;
- 28513 01 - Industrial maintenance paints, interior;
- 28513 05 - Industrial maintenance paints, exterior; and
- 28513 11 - Traffic paints.

In addition, the ARB Methods Manual states that this category accounts for unspecified industrial coatings emissions that are not inventoried in any point source category. Because of a lack of consistent point source data for this category (data are available for every 3 years over the period 1987-1997, but data values jump dramatically between available years), no attempt was made to adjust the Bureau of the Census data to eliminate point source emissions activity.

Pechan compiled 1985-1998 National production data for the Bureau of the Census paint categories used by ARB in developing base year emissions for this source category. In 1994, the Bureau of the Census revised the coding scheme for the first two categories listed above. Because it was clear that three new codes replaced the two previous codes, it was possible to compile a consistent series from 1985-1998. Next, Pechan regressed these National production

data against the following variables: output and employment in the Durables Manufacturing sector and output and employment in the Total Manufacturing sector. The multiple regression analysis indicated that output in the Durables Manufacturing sector provides the best correlation with production of these industrial coatings (adjusted $r^2 = 0.931$, t-statistic of 13.25). Pechan utilized REMI's Durables Manufacturing output data in the equation identified from the analysis that regressed the National industrial coatings production data against these data.

A review of the resulting growth surrogate estimates indicated anomalously large values for several years over the study period, including the years incorporated into the regression analysis. Because of this concern and the fact that California-specific data and area source-specific data were not available for use in the regression analysis, Pechan used county-level trends in the REMI output data for the Durables Manufacturing sector to represent the growth surrogate data for this source category.

Category 03. CES 83030 - Organic Solvent Adhesives and Sealants

To develop base year emissions activity data for this category, ARB's Methods Manual describes the use of data provided in Decision Resources' "Outlook for the U.S. Adhesives and Sealants Industry," February 1981. This article presented adhesives and sealants National production data for various end uses (ARB assumed that production was equal to use). Because ARB assumed that both packaging industry use and wood converting (forest product) industry use are accounted for in the point source inventory, ARB only used the National production data for construction, transportation, and "all other" industries for this category. For the base year inventory, 10 percent of adhesives were assumed to be solvent-based (45 percent were assumed to be water-based, 25 percent hot melts, and 20 percent other). Based on a 1982 SAIC report, "Development and Improvement of Organic Compound Emissions Inventory for California," ARB assumed that California consumed 12.3 percent of the National production of adhesives and sealants. The Statewide emissions estimate was then distributed to counties based on new nonresidential and residential building construction data from Security Pacific Bank's "California Construction Trends," December 1983.

Because ARB developed base year emissions estimates using references that only provide data for a single year, Pechan investigated the availability of other sources of activity data. None of the identified sources provided long-term data specific to solvent-based adhesives and sealants. For example, the *Census of Manufacturers* presents raw material use on a dry-weight basis, which excludes solvents. To represent National emissions activity data for this category, therefore, Pechan developed constant dollar shipments data for SIC code 2891 (Adhesives and Sealants) from the National Bureau of Economic Research-Center for Economic Studies (NBER-CES)'s "Census Manufacturing Industry Productivity Data Base" (NBER, 2000). This data base contains National value of shipments data in nominal dollars and shipments price deflators for this industry for 1970-1994. To develop emissions activity data for 1995-1997, Pechan compiled *Annual Survey of Manufacturers* value of shipments data, which is the source of the data that underlie the NBER-CES data base, and the producer price index for SIC code 2891 (Census, 1997c). These data were then normalized to the NBER-CES data to develop a complete 1970-1997 data series. Pechan then conducted analyses regressing these data against National employment and output data for the following sectors: Construction; Nondurables

Manufacturing; Durables Manufacturing; Total Manufacturing; and Carpets and Rugs. These sectors are the major adhesive and sealants consuming industries described in the *Rauch Guide to Adhesives and Sealants Industry, 1995-1996* or in information obtained from the Freedonia Group (FMC, 1996 and Freedonia, 1999a). Based on these regressions, the indicator that best correlated with the production data was employment in the Construction industry (with an adjusted r^2 of 0.922 and t-statistic of 17.86). Pechan therefore used REMI employment data for the Construction industry in the equation identified from the regression analysis (see Table III-1) to develop trends in adhesives and sealants demand.

Because this method only projects total adhesives and sealants demand, and industry sources indicate declines in solvent-based versus alternative adhesive technologies (e.g., water-borne and hot melts), an adjustment factor was applied to reflect available information on the reduction in solvent-based adhesives and sealants demand in the United States (California-specific data are not available). Based on the percentage of total adhesives demand that is solvent-based that is available from Freedonia for 1989, 1993, 1998, 2003, and 2008, Pechan estimated the percentage of total demand attributed to solvent-based for other years over the period 1970-2030. (Note that industry sources state that 95 percent of total adhesives and sealants demand is attributable to adhesives.) Because of a lack of information, Pechan applied the 1989 percentage to all pre-1989 years. For all other years, Pechan developed a log-linear time series regression fitted to the solvent-based percentage values for 1989, 1993, 1998, 2003, and 2008. The log-linear regression indicated a statistically significant relationship between the year and solvent-based demand as a percentage of total adhesives demand (adjusted $r^2 = 0.996$; t-statistic of -30.84). The resulting regression equation was used to interpolate between the available years from the Freedonia source and to project the percentage of total adhesives and sealants demand that is solvent-based. In 1989, this percentage was 16.1; in 2008, Freedonia projects the percentage at 7.7; in 2030, Pechan forecasts the percentage at 3.3 percent.

The growth surrogate data for this source category was estimated using the regression output described above, adjusted by factors reflecting the reduction in total adhesives and sealants demand attributable to organic solvent-based adhesives and sealants.

Category 04. CES 46441 - Oil and Gas Production, Other, Gaseous Fuel (Unspecified)

The ARB's emissions activity data for this category is the amount of natural gas produced in each county. These data, which are originally reported in an annual report published by the DO&G, "(73rd) Annual Report of the State Oil and Gas Supervisor," were compiled from the "California Statistical Abstract" for 1970-1998 (DOF, 1999). Data were not available from this source for 1979. To estimate data for this year, Pechan interpolated between the values for 1978 and 1980.

Pechan also obtained natural gas supply projections from the CEC for Southern and Northern California for 2002, 2007, 2012, 2017, and 2022 (CEC, 2000). After regressing the CEC forecast data against various REMI variables, Pechan identified a statistically significant equation relating total California on-shore natural gas production to two variables: (1) employment in the Gas Utilities sector; and (2) output in the Crude Petroleum, Natural Gas,

and Gas Liquids sector (adjusted $r^2 = 0.986$; t-statistics of -6.183 and 4.855, respectively). The REMI “best estimate” forecast data for these sectors were input into this equation to yield post-2020 state-level production forecasts and to interpolate between the available CEC forecast year values. Northern and Southern California on-shore production estimates were developed by applying regional ratios to the State totals. To develop regional ratios for forecast years not reported by the CEC, Pechan interpolated between the values for the years forecast by the CEC. To represent post-2022 year ratios, Pechan used the regional ratios for 2022. Because the CEC projections data do not exactly match the CEC historical data, Pechan developed post-1998 natural gas production estimates by applying growth factors to the actual 1998 county-level production data. These growth factors are based on the CEC regional forecast data and the regional forecast data developed by Pechan as described above.

Category 05. CES 46458 - Oil and Gas Production, Crude Oil Production-Tanks

For this category, ARB’s emissions are estimated based on a series of parameters, including storage tank throughput, type of tank, tank diameter, turnover of the liquid in tank, type of product in tank, etc. To develop the base year values for these parameters, ARB used the results of a one-time oil production tank survey. Pechan was unable to identify historical time-series data for most of the parameters included in the base year emissions estimate. However, county-level oil production data are available from the same DO&G source described above for CES 46411. Pechan compiled these 1970-1998 county-level oil production data to represent the historical emissions activity data for this category. Because data are not available for 1979, Pechan interpolated the values for this year. To project oil production, Pechan conducted analyses regressing the historical state-level oil production data with various indicators available from REMI’s “best estimate” forecast models. The regression analyses identified the following variables as best correlated with California oil production: (1) employment in the Crude Petroleum, Natural Gas, and Gas Liquids sector and (2) employment in the Pipelines, excluding Natural Gas sector (adjusted $r^2 = 0.878$; t-statistics of 14.03 and -8.408, respectively). Pechan developed post-1998 forecasts of county-level oil production by applying growth factors to the actual 1998 county-level oil production. These growth factors were developed based on output from the equation described above and relevant employment data from REMI’s “best estimate” forecasting models.

Category 06. CES 83089 - Consumer Products, Non-Aerosol Solvents

According to ARB’s Methods Manual, a diverse set of products comprise this category, which includes products used in residential, institutional, and commercial establishments. Products include, but are not limited to: window cleaners, general purpose cleaners, spot removers, floor and furniture polishes, household adhesives and sealants, electric pre-shave, non-aerosol hair care products, after shave, stick and roll-on deodorants, nail care products, rubbing alcohol, mouthwash, lotions, radiator and windshield washer antifreezes, and automotive brake fluid. Table III in Section 6.1 of ARB’s Methods Manual lists the estimated statewide solvent use for non-aerosol consumer products for 1987. The following lists the products from this table with solvent use of more than 1,000 tons per year:

1) Radiator antifreeze	117,860*
2) Non-aerosol windshield washer antifreeze	10,450
3) Non-aerosol hair sprays	2,582
4) Floor polish	2,336
5) Stick deodorants, anti-perspirants	1,857
6) Non-aerosol after shave	1,542
7) Creams	1,490
8) Brake fluid	1,485
9) Household adhesives and sealants	1,442
10) Furniture polish	1,422
11) Rubbing alcohol	1,316
12) Mouthwash	1,118
13) Shampoo	1,021

Notes: Emissions for boldface products are based on national 1972 *Census of Manufacturers* data.

* Only 5 percent of this amount (5,893) was assumed to actually be emitted into air.

Using ARB's assumption that 5 percent of radiator antifreeze is emitted into the atmosphere, the top two consumer products emitted 42 percent of the 38,892 total solvents emitted from this category. The top 5 and top 10 products emitted 59 percent and 78 percent, respectively, of the total consumer product solvents emitted. ARB used numerous references in compiling base year emissions activity for this source category. However, the 1972 *Census of Manufacturers* was the source of these data for most of the individual products representing greater than 1,000 tons per year of solvent use (Census, 1972).

Although the "Soaps, Cleaners, and Toilet Goods" industry report from the *Census of Manufacturers* presents National production data for most of the top 10 products in this category, it is important to note that this was not ARB's data source for automotive washer fluid use. Activity data for this consumer product was based on a one-time survey conducted by Charles H. Kline and Company. Pechan attempted to develop a long-term production series for the major contributors to this source category as identified in ARB's Methods Manual. Unfortunately, *Census of Manufacturers* data are only available every 5 years. In addition, the coverage of consumer products was inconsistent across years, with the most glaring omission being the lack of pre-1992 data for the single largest emissions contributor, windshield washer antifreeze. Pechan was also unable to identify long-term data for the consumer products for which ARB estimated emissions based on non-*Census of Manufacturers* references.

Because of these problems, Pechan researched the availability of other sources of consumer products data. Pechan identified that Freedonia's "Solvents to 2003" and "Automotive Fluids and Chemicals to 2003" publications contain 1989, 1993, 1998, 2003, and 2008 solvents data for: (1) consumer products markets; and (2) antifreezes and deicers (Freedonia, 1999b and Freedonia, 1999c). Pechan compiled these National consumer product solvents demand data from Freedonia. Solvent demand data are available from this source for the following specific consumer products categories:

- Household cleaners;
- Toiletries, cosmetics, and drugs;

- Dry cleaning;
- Antifreeze and deicers; and
- Other applications.

Because ARB categorizes emissions from solvents used for dry cleaning in separate ARB emissions categories (e.g., CES 46797), these data were not included in this analysis. Following ARB's emission estimation methods, Pechan estimated that 13 percent of the antifreezes/deicers solvent use is emitted into the air. This 13 percent was computed using 1987 estimates of radiator and antifreeze consumption and assumptions that all windshield antifreeze, but only 5 percent of radiator antifreeze is actually emitted into the air. No other adjustments were made to the Freedonia estimates because ARB assumes that 100 percent of the solvents in all other consumer products are ultimately emitted into the air.

For many of the consumer products that comprise this category, ARB used population data to allocate National production data to California. For other consumer products (e.g., windshield washer antifreeze), ARB used vehicle registration data to allocate National data to California. Because Pechan does not have access to vehicle registration forecasts, Pechan analyzed VMT data as a potential explanatory variable for consumer product solvent use. Pechan developed a multiple regression analysis of the Freedonia solvents demand data against National data for the following variables: population; employment and output in the Motor Vehicles and Equipment sector; real disposable personal income; and VMT. The equation identified from the analysis included population as strongly correlated with the emissions activity data (adjusted $r^2 = 0.990$, t-statistic of 20.16).

To estimate 1970-2030 growth surrogate data for this source category, Pechan used the identified equation relating solvent use to population data. Pechan reviewed these data for reasonableness and identified large growth rates that appear unsustainable throughout the entire study period. Therefore, Pechan revised the growth surrogate data for the years before and after the period included in the regression analysis (i.e., 1989-2008). To represent pre-1989 and post-2008 trends in this category, Pechan used REMI's county-level population growth rates from their "best estimate" forecast scenario.

Category 07. CES 83550 - Agricultural Pesticides, Methyl Bromide

ARB provided Pechan with a file developed by the DPR (Lerch, 2000a). This file contains 1990-1998 process rates and total organic gas and ROG emissions for all four agricultural/structural pesticide categories in section 6.4 of ARB's Methods Manual at the Air Basin/District/county level. The DPR has only required full reporting of pesticide use since 1990; since 1970, DPR has compiled data from all pesticides used by commercial pest control operators and restricted material applications from farmers (i.e., criteria for restricted material designation includes harm to public health, farm workers, and others). DPR has published annual reports on these data since 1971. These tabular reports include the pounds applied and the number of acres or other units treated. Pechan compiled the pre-1990 DPR data, but determined that these data should not be used as surrogate growth data because they are not comparable to data in later years due to the significant change in reporting requirements and because of some anomalous changes in year-to-year values. (It should be noted that the DPR

suggested that there may be gross errors in these data as they did not quality assure the pre-1990 data that were reported to them).

DPR staff has recommended to ARB that they incorporate the following assumptions into their forecasts of the two methyl bromide source categories to reflect the impact of the Montreal Protocol: (1) for Agricultural Pesticides-Methyl Bromide (CES 83550), the source category should grow through 2001, then be reduced to one-third of the 1990 value by year 2010; and (2) for Structural Pesticides-Methyl Bromide (CES 83576), the source category should be reduced to zero by the year 2005.

Pechan developed regression analyses of the 1990-1998 DPR data for this source category with a dozen variables (e.g., output and employment in farm sector; output and employment in food manufacturing sector), but was unable to identify a successful correlation with any of these variables. Based on consultation with the ARB Contract Manager, Pechan used county-level Farm output as the surrogate growth indicator for pre-1990 trends in this category. Output in the Farm sector was used because of the strong historical correlation between agricultural non-methyl bromide use and this indicator. For 1990-1998, Pechan used actual Air Basin/District/county estimates of the amount of methyl bromide applied as reported by the California DPR. Pechan forecasted this emissions activity based on the DPR's assumptions. Specifically, Pechan grew the 1998 DPR estimates through 2001 using the county-level change in Farm sector output obtained from REMI's "best estimate" forecast. Pechan then estimated post-2001 values, by reducing the 2001 estimates so that by the year 2010, state-level agricultural methyl bromide is reported as one-third of the 1990 value. Pechan assumed no change in agricultural methyl bromide use after 2010. Pechan developed pre-1990 estimates by applying the county-level trends in Farm output to the 1990 DPR data.

[Note that ARB also provided Pechan with 1990-1998 DPR data on structural use of methyl bromide. ARB stated that Pechan should assume that structural use of methyl bromide (CES 83576) should be reduced to zero by year 2005. Pechan did not try to collect pre-1990 data for this category as it is not one of the categories identified for detailed review. To back-cast this category, however, Pechan developed a regression equation based on the state-level 1990-1998 DPR data. Nonresidential fixed investment and population were identified as statistically correlated with this emissions activity (adjusted $r^2 = 0.979$, t -statistics of 10.68 and -10.27, respectively). However, this equation resulted in unreasonably large pre-1990 values. The reason for this result is the uncharacteristically large decline reported between 1990 and 1998 in structural use of methyl bromide. The two explanations for this dramatic decline are the tightened pesticide use reporting requirements that went into effect in 1990, and the impact of the Montreal Protocol. Because of the data anomalies resulting from the implementation of the regression-based approach, Pechan developed pre-1990 growth surrogate data for this category based on the county-level trend in non-residential fixed investment as reported by REMI. Pechan used this variable because the regression analysis indicated that it was the most strongly correlated with the DPR emissions activity data. Pechan developed post-1998 activity for this category by reducing the 1998 DPR estimates to zero in 2005 using a constant rate of decline.]

Category 08. CES 83568 - Agricultural Pesticides, Non-Methyl Bromide

As described above for CES 83550, Pechan obtained DPR non-methyl bromide agricultural use data for 1990-1998 from ARB and compiled 1970-1989 use data from the DPR. The pre-1990 DPR estimates are very different from the estimates from 1990 due to reporting requirement changes (improvements in pesticide use reporting coverage). Because of this inconsistency, Pechan did not use the pre-1990 DPR data in this effort. Based on a regression analysis, Pechan identified REMI Farm sector output data as statistically correlated with the application of non-methyl bromide for agricultural use (adjusted $r^2 = 0.916$, t-statistic of 9.395). Pechan used the DPR Air Basin/District/county pesticide use data for the 1990-1998 growth surrogate data for this category. Pechan used county-level REMI Farm sector output data in the equation identified from the regression analysis in back-casting and forecasting growth surrogate data for this category.

[Note that ARB also provided Pechan with 1990-1998 DPR data on structural use of non-methyl bromide pesticides. Pechan did not try to collect pre-1990 data for this category as it is not one of the 50 categories identified for detailed analysis; Pechan developed a regression equation based on the state-level 1990-1998 DPR data. Housing expenditures were identified as statistically correlated with this emissions activity (adjusted $r^2 = 0.658$, t-statistic of 4.051). Pechan used REMI estimates for this variable and the identified equation to estimate growth surrogate data for the category. However, the resulting data provided numerous unreasonably large values. Because of these data anomalies and due to the short time frame and relatively weak correlation identified from the regression analysis, Pechan did not use the regression-based approach to back-cast/forecast this category's growth surrogate data. Pechan instead directly used the county-level trend in housing expenditures as reported in REMI's "best estimate" model to develop pre-1990 and post-1998 estimates for this category.]

Category 09. CES 82123 - Residential Wood Combustion, Fireplaces

ARB's emissions estimation method for this category is to multiply an average wood consumption per fireplace estimate (0.28 cords per year) by the estimated number of houses with active fireplaces. ARB derived the estimated number of houses with active fireplaces by subtracting the estimated number of wood heating houses (available from the 1990 *Census of Population and Housing-Summary Tape File 3A*) from the total number of wood burning houses. The total number of wood burning houses was estimated by multiplying the fraction of houses burning wood by the total number of houses. Table III in Northern California Research Associates' 1988 report "The California Residential Wood Consumption Survey," lists the estimated percentage of wood burning households by Air Basin.

Because the references used by ARB in developing base year emissions for this category do not report similar data for other years, Pechan compiled data from the Bureau of the Census and U.S. Department of Housing and Urban Development's "American Housing Survey for the United States" for 1987, 1989, 1991, 1993, 1995, and 1997 on the number of fireplaces without inserts in the West Region of the U.S. (HUD, 1997). (It should be noted that fireplaces with inserts are included in ARB's wood stove source category.) From this same source, Pechan also compiled data for these same years on the total number of occupied housing units in the West

Region and then developed the percentage of total occupied housing units that have fireplaces without inserts in the West Region for each year. Unfortunately, the 1997 percentages cannot be used as the data are not consistent with previous years due to changes to the questionnaire used:

PERCENTAGE OF HOUSING UNITS WITH FIREPLACES WITHOUT INSERTS	
YEAR	
1987	19.4
1989	18.1
1991	17.3
1993	15.6
1995	14.4
1997	6.5

It should be noted that these percentages are based on the total number of fireplaces without inserts, which includes some natural gas fireplaces. Based on DOE's "Housing Characteristics" publications for 1993 and 1997, 86.5 percent of 1993 households with fireplaces used as secondary heating equipment in West Region used wood, with the remaining fireplaces using natural gas (DOE, 1997). The wood-fired percentage dropped to 84.4 percent of households in 1997. Because of the high percentage of total fireplaces without inserts that use wood, Pechan assumes that these data can be used in developing trends in wood burning fireplace activity.

Pechan also compiled data on the number of occupied housing units in each California county for 1980-1999 from the California Department of Finance (DOF, 1992 and DOF, 2000). These data were used in a multiple regression analysis to estimate the number of occupied housing units in California back through 1970 and forward through 2030. The analysis regressed the state-level number of occupied housing units against the following REMI variables: population, housing expenditures, and residential fixed investment. The regression analysis identified population data as strongly correlated with the number of occupied housing units (adjusted $r^2 = 0.996$; t-statistic of 67.42). The resulting regression equation was used with REMI population data to estimate occupied housing units growth surrogate data for 1970-2030.

The occupied housing units data were then multiplied by a factor representing the trend in the percentage of households that have fireplaces without inserts (based on the DOE West region data). Because percentages were only available for odd years between 1987 and 1995, Pechan interpolated the percentages for even years during this period. Pechan then applied a factor to each year's growth surrogate data to adjust for the known trend in housing units with fireplaces without inserts. Specifically, Pechan applied a factor of 1 to every 1970-1987 record, and a factor of less than one for every post-1987 year. The post-1987 factors were based on the ratio of the current year's percentage to the percentage in 1987. For example, a factor of .93 was applied to all 1989 records ($.93 = 18.1/19.4$). The 1995 factor was applied to all post-1995 records due to the lack of information about trends after that year.

Category 10. CES 54569 - Residential Fuel Combustion-Space Heating, Natural Gas

The emissions activity for this category is millions of cubic feet of residential natural gas consumed for space heating. ARB developed base year emissions data for this category by multiplying county-level residential natural gas consumption from the CEC by fractions representing the portion of total residential natural gas used for space heating (region-specific fractions were determined by ARB).

For this effort, Pechan obtained a file from the CEC providing residential “central space heating” natural gas demand for 1970-2017 for each of three gas service territories – Pacific Gas and Electric (PGE), Southern California Gas (SCG), and San Diego Gas and Electric (SDGE), and for 1980-2017, for an “other” territory designation (Gough, 2000). The “other” gas service territory, which represented less than 1 percent of total State natural gas demand in 1997, includes demand centers that are located in counties adjacent to the California-Oregon and California-Nevada borders, and a portion of San Bernardino county.

To develop growth surrogate data for this category, Pechan conducted multiple regression analyses with the state-level 1970-2017 residential “central space heating” natural gas demand data. The potential explanatory variables investigated in this analysis included: output and employment in the Real Estate sector; population; real disposable income; housing expenditures; and fixed residential investment. Based on the regression analysis, Pechan identified a strong inverse correlation (adjusted $r^2 = 0.936$, t-statistic of -26.15) between residential natural gas space heating demand and population (i.e., as population increases, the demand for natural gas for space heating declines). This result can be explained by increases in residential space heating energy efficiency. Pechan used the established regression equation and REMI population data to develop county-level residential space heating natural gas consumption throughout the analysis period.

Category 11. CES 54577 - Residential Fuel Combustion-Water Heating, Natural Gas

Millions of cubic feet of residential natural gas consumed for water heating is the emissions activity data for this source category. To develop base year emissions estimates, ARB multiplied total county residential natural gas consumption by region-specific fractions representing the percentage of total consumption used for water heating. The region-specific fractions were determined by ARB.

Pechan obtained a file developed by the CEC that provides residential “indirect hot water heating for clothes washing,” “indirect hot water heating for dishwashing,” “pool heating,” “hot tub fuel,” and “water heating” natural gas demand for 1970-2017 for each of three gas service territories (PGE, SCG, SDGE), and for 1980-2017, for an “other” territory designation (Gough, 2000). The “other” gas service territory, which represented less than 1 percent of total State natural gas demand in 1997, includes demand centers that are located in counties adjacent to the California-Oregon and California-Nevada borders, and a portion of San Bernardino county.

To develop growth surrogate data for this category, Pechan conducted an analogous multiple regression analysis to that described above for CES 54569. The analysis indicated that real disposable income is strongly correlated with CEC's total residential water heating natural gas demand data. Pechan used the identified equation and county-level forecasts of real disposable income from REMI to estimate the growth surrogate data for this category.

Category 12. CES 47332 - Agricultural Land Preparation

The number of acre-passes used in crop planting/harvesting is the emissions activity data for this category. The following equation identifies how emissions are estimated:

$$\text{Emissions}_{\text{crop}} = \text{Emission Factor} * \text{Acres}_{\text{crop}} * \text{Acre-passes/acre}_{\text{crop}}$$

In developing base year emissions, ARB used harvested crop acreage data from the California Department of Food and Agriculture (CDFA) and ARB estimates of the number of passes per crop.

Pechan developed acre-passes data for 1986-1998 from CDFA crop acreage data and crop profiles identified in ARB's Methods Manual (CDFA, 2000). Because of changes in the crops reported over the period, the first step in this process required that Pechan identify a consistent set of crops for which data are available throughout the 1986-1998 period. After identifying these crops, Pechan computed the total acreage for these crops in each year. In some rare instances, the CDFA did not report harvested acreage data for a small number of years for a county even though that county showed substantial acreage values for all other years throughout the period. For these instances, which are likely caused by confidentiality concerns, Pechan estimated missing year acreage values using linear interpolation. In cases where data were missing for initial years during the 1986-1998 period, Pechan estimated acreage totals using the first available year's acreage value. An analogous procedure was used to estimate values for missing acreage data occurring at the end of the 1986-1998 period. Pechan then compiled ARB's acre-pass profile values for over 150 crop types. These profiles were multiplied by the 1986-1998 county-level harvested crop acreage data to estimate the emissions activity for this source category (total annual acre-passes) over the 1986-1998 period.

Pechan conducted multiple regression analysis with a number of farm-related variables and identified an equation with the best statistical fit against the state-level 1986-1998 annual acre-passes data. This equation includes the following three explanatory variables: employment in the Agricultural Services sector (negative coefficient); employment in the Farm and Garden Machinery sector, and employment in the Farm sector (adjusted $r^2 = .854$; t-statistics of -6.380, 4.200, and 2.663, respectively). Pechan used the identified equation and REMI county-level historical and forecast employment data for these sectors to develop growth surrogate data for this category.

Category 13. CES 47340 - Farming Operations, Cattle Feedlot Dust

The emissions activity for this source category is based on beef cattle feedlot throughput. ARB defines beef cattle feedlot throughput as the number of beef cattle that have been confined

in feedlots for 4 to 5 months for fattening before marketing. For the purposes of estimating emissions, ARB relied on two sets of available data: (1) the state-level annual number of beef cattle marketed from feedlots; and (2) the monthly total number of cattle in feedlots by California District (these data are not equivalent to feedlot throughput). Pechan compiled the annual number of beef cattle marketed from California feedlots for 1970-1998. Data for 1970-1988 were from annual editions of *California Livestock Statistics*, published by the CDFA, while data for 1989-1998 were from the CDFA's *1998 Agricultural Resource Directory* (CDFA, 1988 and CDFA, 1998).

To geographically allocate state-level activity for this category, ARB also compiled data on the monthly number of cattle in feedlots for each of four California Districts. Pechan compiled these data for 1970-1988 from annual editions of *California Livestock Statistics* for the following four districts: Sacramento Valley, North Coast, Central Coast; San Joaquin Valley, less Kern County; Southern California & Kern County less Imperial Valley; and Imperial Valley. This report ceased publication after 1988. ARB averaged the monthly data over a 12-month period to disaggregate the throughput to the region level.

To develop 1970-1988 growth surrogate data for this category, Pechan used a combination of state-level feedlot throughput data and regional total feedlot cattle data. Specifically, Pechan developed regional growth factors based on allocating the state-level throughput data to regions using the total number of feedlot cattle in each year. These growth factors were applied to the 1987 activity data presented in Table III of Section 7.6 of ARB's Methods Manual to estimate the feedlot throughput for each county in a region. To develop 1989-1998 growth surrogate data, Pechan applied the state-level change in feedlot throughput to the estimated 1988 county-level surrogate activity data.

To assist in developing post-1998 emissions activity data for this category, Pechan conducted a regression analysis to identify surrogate indicators that historically correlated with the State number of cattle marketed from feedlots. Based on this analysis, Pechan identified a strong inverse correlation between output (sales) in the Meat Products sector and the number of cattle marketed from feedlots (adjusted $r^2 = 0.922$, t-statistic of -18.19). This indicates that as output in the Meat Products sector has increased historically, the number of cattle marketed from feedlots has declined. Pechan believes that one potential explanation for the historical inverse correlation is the shift in demand from beef to non-beef meat products (especially chicken).

Using the identified equation relating Meat Products sector output to the number of cattle marketed from feedlots and using REMI projections of the Meat product sector output, results in zero cattle marketed from California feedlots by the year 2007. Pechan contacted the California Cattlemen's Association (CCA) for insight into whether this is expected to occur. A CCA feedlot expert stated that the number of feedlots has declined in California since the 1950s because meat packing houses have moved from California to the Midwestern United States. The reason for this shift is the relative increase in California's labor costs. Although he could not speak to the number of cattle in feedlots, the CCA expert noted that there are approximately 15 cattle feedlots left in the State, with nearly all feedlots in the Imperial Valley. The CCA contact noted that one large feedlot currently exists in the Central Valley, and that this feedlot owns its own packing house, and is likely to remain in business. However, the other feedlots in Imperial Valley are

currently struggling to find a packing house and are considering building their own. He noted that if they are not able to do so, they will likely shut down. Given the uncertainty surrounding the future level of emissions activity for this category, Pechan incorporated a no growth assumption for all post-1998 years.

Category 14. CES 47357 - Building Construction Dust, Residential

The ARB emissions activity data for this source category is annual acre-months of residential construction. To develop these data for the base year inventory, ARB compiled the number of single-family and multi-family residential units constructed. The ARB's Methods Manual states that these data were obtained from the California Department of Finance (DOF). For this effort, Pechan compiled 1970-1999 residential permit data from the Construction Industry Research Board (CIRB), which is the original source of the California DOF data cited in ARB's Methods Manual (CIRB, 2000). Although the CIRB data are available by county from 1967-1999, it is important to note that the pre-1980 data are not comparable to the 1980-1999 data (prior to 1980, permits for alterations and additions of \$100,000 or more were counted as new construction). For the base year inventory, ARB converted the number of single-family and multi-family construction permits based on an assumed construction project duration of 6 months for both types of residential construction. ARB also assumed one-twentieth of an acre of construction for each multi-family unit and either one-fifth of an acre or one-seventh of an acre for each single-family unit. ARB assigned the one-fifth of an acre assumption to heavily populated counties (45 counties) and the one-seventh of an acre assumption to less populated counties (13 counties). Pechan used these ARB assumptions to convert the CIRB permit data to annual acre-months of residential construction.

Next, Pechan conducted a multiple regression analysis to identify variables that relate to annual acre-months of residential construction. Because of the change in the CIRB's new construction definition that occurred in 1980, Pechan only included 1980-1999 data in the regression analysis. Pechan regressed the 1980-1999 state-level annual acre-months of residential construction data against the following variables: employment and output in the Construction sector; population; employment and output in the Real Estate sector; and housing expenditures. The equation with the best statistical fit includes both employment and output in the Construction sector ($r^2 = 0.904$; adjusted $r^2 = 0.893$; t-statistics of -7.091 and 12.18, respectively).

Pechan used the Construction sector employment and output data and the identified equation to estimate growth surrogate data for this category. A review of the resulting data identified large growth rates over the forecast period. Pechan, therefore, developed final growth surrogate data for this category by using the regression-based values only for the years included in the regression analysis (1980-1999). For all other years, Pechan applied the trend in output in the Construction sector. This approach was used because Construction sector output had the strongest correlation with the historical emissions activity data for this category as identified from the regression analysis.

Category 15. CES 47365 - Building Construction Dust, Commercial

The emissions activity for CES 47365 is acre-months per year of commercial construction. ARB used commercial construction valuation data published by the California DOF in estimating base year emissions activity. For this effort, Pechan compiled available time-series commercial construction valuation data from the CIRB, which is the original source of the California DOF data cited in ARB's Methods Manual (CIRB, 2000).

For the base year inventory, ARB converted the total dollar of commercial building construction into acre-months per year based on an assumed 3.7 acres per million dollars of commercial construction valuation and an assumed construction project duration of 11 months. ARB first converted the dollar value of commercial construction from nominal dollars to real dollars using composite construction cost indices published by the U.S. Department of Commerce's Bureau of the Census. Pechan converted the county-level CIRB 1980-1999 commercial construction data to real dollars using the Bureau of the Census' fixed-weighted construction cost index.

As noted in the discussion above for CES 47357, the pre-1980 CIRB data are not comparable to the 1980-1999 data due to a change in the definition of new construction. After converting the CIRB 1980-1999 state-level commercial construction valuation data to real dollars using the U.S. Bureau of the Census' fixed-weighted construction cost index, Pechan converted these data to acre-months per year using ARB's base year emissions estimation assumptions and procedures. Pechan then developed a multiple regression analysis using this state-level data series and REMI state-level data for population and employment and output in the following sectors: Construction; Real Estate; and Commercial (a combination of the Finance, Insurance, and Real Estate; Retail and Wholesale Trade; and Services sectors). The equation with the best statistical fit includes both employment and output in the Construction sector as explanatory variables (adjusted $r^2 = 0.863$; t-statistics of 10.73 for construction output and -9.806 for construction employment). Pechan developed output from the identified equation after inputting REMI Construction sector output and employment data. A review of the resulting data indicated very large growth rates over the forecast period. Because Pechan deemed these growth rates as suspect, Pechan developed the final surrogate data for this category based on a combination of regression equation output (for 1980-1999) and output data for the Construction sector, which was the variable identified as most strongly correlated with the emissions activity data for this category for (for all other years).

Category 16. CES 47381 - Road Construction Dust

The ARB Methods Manual notes that the emissions activity data for this category is based on estimates of the number of miles of road built and the assumed acreage disturbed per mile of construction. The ARB estimated the number of miles built in the base year by taking the difference in total road miles between consecutive years. The ARB assumed that the number of acres disturbed varies for three road types: freeways, State highways, and city and county roads. The ARB cites the California DOF and the California Department of Transportation (Caltrans) as sources of the data on the miles of new road built.

Pechan compiled 1980-1998 data from Caltrans on the mileage of maintained roads by county and jurisdiction (Caltrans, 1990). After reviewing the 1980-1998 data series, Pechan contacted Caltrans to determine why there is a negative change in the miles of road maintained for numerous years within this time frame. Discussions with Caltrans failed to provide Pechan with an approach for overcoming this problem. Because of this shortcoming, Pechan will use ARB's current growth surrogate (i.e., employment in the Construction sector—SIC codes 15 through 17) to represent emissions activity for this category. Because REMI has developed forecasts for this variable under four different growth scenarios (i.e., best estimate, low, high, and cyclic), these alternative forecasts are available for use as growth surrogate data for this category.

Category 17. CES 47399 - Unpaved Road Travel Dust-City and County Roads

The ARB Methods Manual states that base year emissions for this category are based on the number of unpaved city and county road miles and an assumed 10 daily VMT per road mile. Pechan compiled 1981-1992 state-level data on unpaved city and county road mileage, which is available from Caltrans' "Assembly of Statistical Reports" for 1981-1992 (Caltrans, 1998). The Caltrans post-1992 data report total road mileage for city and county roads; since most city and county roads are paved, trends in these data were not used as a surrogate for trends in unpaved city and county road mileage. County-level unpaved road mileage data are not available for unpaved city and county roads (only total unpaved road data are reported by county).

Pechan conducted a multiple regression analysis to identify potential surrogates that correlate with the state-level 1981-1992 road mileage data. One of the potential explanatory variables included in this analysis was 1970-2020 light- and medium-duty truck (LDT and MDT) VMT data provided by the ARB Contract Manager. Based on the analysis, Pechan identified total LDT and MDT VMT data to be strongly correlated with unpaved city and county road mileage (adjusted $r^2 = 0.932$; t-statistic of -12.30). The analysis indicates an inverse correlation between the two variables (i.e., as LDT/MDT VMT increases, city and county unpaved road VMT declines). The likely explanation for this phenomenon is an increase in the relative proportion of LDT/MDT use on paved roads as there has been a substantial increase in the purchase and use of pick-up and sport-utility vehicles on paved roads. Pechan used the Air Basin/county-level LDT/MDT VMT data and the identified equation relating these data to unpaved city and county road VMT to estimate the trends in unpaved city and county road VMT from 1970 through 2030 (after extrapolating the available LDT/MDT VMT data to 2030). Pechan's review of the resulting estimates indicated near zero unpaved road mileage by the year 2030. In lieu of any information that this estimate is reasonable, Pechan employed the regression output as the growth surrogate data only for the years for which the regression analysis was conducted (i.e., 1981-1992). Post-1992 values were based on a no growth assumption since the strongest correlated variable (LDT and MDT VMT) indicated an inverse relationship; this assumption was also employed for all pre-1981 years.

Category 18. CES 47423 - Unpaved Road Travel Dust-Bureau of Land Management and Bureau of Indian Affairs Roads

The ARB Methods Manual states that base year emissions for this category were based on the number of unpaved Bureau of Land Management (BLM) and Bureau of Indian Affairs (BIA) road miles and an assumed 10 daily vehicle miles traveled per mile of road.

Pechan compiled state-level data on unpaved road mileage for BLM and BIA roads, which is available from Caltrans for 1981-1992 (county-level data for BLM and BIA unpaved roads are not available). To estimate 1993-1996 BLM and BIA unpaved road miles, Pechan applied the rate-of-change in total BLM and BIA road mileage to the 1992 unpaved road mileage. This approach assumes that the trend in BLM and BIA total road mileage is the same as the trend in unpaved BLM and BIA road mileage. This assumption is supported by the fact that the vast majority of BLM/BIA roads are unpaved (e.g., approximately 97 percent in 1992).

Pechan conducted regression analyses to identify surrogates that correlate with the available 1981-1996 BLM and BIA unpaved road mileage data. The best statistical fit was identified with two surrogates: Population and the number of forestry and logging occupations (adjusted $r^2 = .761$, t-statistics of 7.039 and -2.505, respectively). Unlike the population variable, the forestry and logging occupations variable is inversely correlated to the emissions activity. This may result from a relative increase in demand for these roads for recreational use versus logging use. To develop growth surrogate data for this category, Pechan first used the REMI population and forestry and logging occupations data in the established equation described above. After identifying some data anomalies concerning large growth rates for this category, Pechan decided to employ a combination of regression output (for years covered by regression analysis) and trends in population for all other years as the growth surrogate data for this category. Pechan selected population data for this effort because the regression analysis indicated that these data were the most strongly correlated with the emissions activity data for this category.

Category 19. CES 47431 - Unpaved Road Travel Dust-Farm Roads

As described in ARB's Methods Manual, the base year emissions for this category are based on an assumed 175 VMT per 40 harvested acres of total non-pasture crops. The ARB obtained total non-pasture harvested crop acreage data from the CDFA.

Pechan compiled 1986-1998 county-level total non-pasture harvested acreage data from the CDFA (see the discussion for category 12 above for details on this compilation). To identify potential growth surrogates for this category, Pechan regressed the state-level harvested acreage data against multiple variables, including employment and output in the following sectors: Meat Products; Dairy Products; Farm and Garden Machinery, Food Manufacturing, Agricultural Services; and Farm. Pechan also included population as a potential independent variable in the regression analysis. Based on this analysis, Pechan identified two variables providing the best statistical fit: population and employment in the Farm and Garden Machinery sector (adjusted $r^2 = 0.793$, t-statistics of -5.098 and 4.111, respectively). The negative coefficient for the population variable can be explained by crop acreage losses as urbanization takes place.

Pechan prepared the growth surrogate data for this category from the identified equation that related population and employment in the Farm and Garden Machinery sector to total non-pasture harvested acreage data.

Category 20. CES 83337 - Fugitive Windblown Dust from Agricultural Lands (Non-Pasture)

ARB calculated base year emissions for this category using a wind erosion equation (WEQ) that develops emission estimates by month. There are many factors that go into this equation, including several factors that are crop-specific. The documentation for the emissions estimation procedure for this category does not describe the values for all of these factors. Even if these data were available, project resources would not be sufficient to calculate emissions by crop by month for each county. The ARB method uses total non-pasture harvested acreage excluding acreage from orchards and vineyards as a key input to the wind erosion equation. The ARB Contract Manager agreed that these data would be used to estimate emissions activity for this source category.

Pechan compiled 1986-1998 data from the CDFA representing total non-pasture harvested acreage excluding orchards and vineyards harvested acreage (see the discussion for category 12 above for additional information on this compilation). Pechan performed a multiple regression analysis to identify variables that correlate with the state-level 1986-1998 acreage data. Potential explanatory variables included in this analysis were: output and employment in the Farm sector; output and employment in the Farm and Garden Machinery Equipment sector; output and employment in the Food Manufacturing sector; output and employment in the Agricultural Services sector; and population. Based on the multiple regression analysis, Pechan identified the best statistical fit with population and employment in the Farm and Garden Machinery Equipment sector (adjusted $r^2 = 0.843$, t-statistics of -6.660 and 3.937, respectively). The potential explanation for the inverse correlation between population and non-pasture acreage is increased urbanization leading to conversion of agricultural land to non-agricultural uses.

To develop growth surrogate data for this category, Pechan compiled the 1970-2030 output of the equation described in the above paragraph after inputting REMI data for population and for employment in the Farm and Garden Machinery Equipment sector.

Category 21. CES 83352 - Fugitive Windblown Dust from Unpaved Roads and Associated Areas

The ARB emissions activity for this category is acreage of unpaved roads. ARB's Methods Manual describes estimating unpaved road acreage in the base year by multiplying total unpaved road mileage by an average unpaved road width of 20 feet. The unpaved road mileage data are from Caltrans' "Assembly of Statistical Reports."

Pechan compiled 1985-1992 total unpaved road mileage data by county from Caltrans' "Assembly of Statistical Reports" (Caltrans, 1998). Because county-level unpaved road mileage data were not available after 1992, Pechan estimated unpaved road mileage for 1993-1996 by multiplying the 1992 county unpaved road mileage data by the state-level ratio of unpaved road

mileage in each year to the 1992 unpaved road mileage. A review of the Caltrans data indicated a dramatic one-time drop in unpaved road mileage between 1986 and 1987. Because the pre-1987 data were nearly double the values for 1987-1996, Pechan excluded the 1985 and 1986 data from the regression analysis. Pechan conducted regression analyses from the 1987-1996 state-level unpaved road mileage data and determined that LDT/MDT VMT and population data provide the best statistical relationship with the historical unpaved road acreage data (adjusted $r^2 = 0.911$, t-statistics of -5.230 and 4.341, respectively). The likely explanation for the negative coefficient associated with LDT/MDT VMT data is an increase in the relative proportion of pick-up truck and sport utility vehicle use on paved roads as there has been a substantial increase in the purchase and use of these vehicles over the historical period.

To develop growth surrogate data for this category, Pechan first developed output from the equation identified above based on REMI population estimates and ARB LDT/MDT VMT estimates (extrapolated to 2030 by Pechan). This output presented zero state-level values by the year 2012. Because it is not clear that unpaved road mileage will decline to zero in the future, and because the variable with the strongest correlation to unpaved road mileage (LDT/MDT VMT) has an inverse relationship with the emissions activity data, Pechan implemented a no growth assumption for all post-regression years for this category. This assumption was also applied to all pre-regression years. Note that this approach is similar to that applied to category 17, which is a very similar source category.

Category 22. SCC 10100601; SIC code 4931 - Electric Generation, Natural Gas, Boilers > 100 MMBtu except Tangential

This is a point source category that is not included in ARB's Methods Manual. Based on the SCC, the emissions activity for this category is the amount of natural gas burned. Pechan first compiled Air Basin/District/county electric utility natural gas consumption projections from the CEC (Lerch, 2000b). Pechan compared these data to the complete list of Air Basin/District/county combinations in California. Pechan developed data for all missing combinations for which Air Basin-level data were available from the CEC source. Specifically, Pechan compiled data for these Air Basin/District/county combinations by assigning Air Basin-level data to these records from the appropriate Air Basin. As noted in Section III.D., ARB requested that Pechan use the CEC data as growth surrogates for all point sources with electric utility-related SIC codes (i.e., 4911 and 4931).

To represent pre-2000 data for this category, Pechan applied state-level growth rates to the CEC year 2000 data. For 1976-1998 these growth rates were compiled from CEC electric utility natural gas consumption estimates (Ewing, 2000). The 1999 rate was estimated based on the average historical growth rate over the 1976-1998 period (approximately 3.5 percent). The 1970-1975 growth rates were compiled from a DOE source of state-level energy consumption estimates (DOE, 2000).

To assist in forecasting growth surrogate data for 2021-2030, Pechan conducted a regression analysis of REMI variables to identify surrogates that correlated with the CEC's 2000-2020 data. The resulting equation identified output in the Electric Utility sector as strongly correlated with the CEC data (adjusted $r^2 = 0.984$; t-statistic of 35.63). Pechan compiled the 2020-2030 county-

level output from this equation based on REMI “best estimate” forecasts of Electric Utility sector output. Pechan then applied the growth rates from the regression output to the CEC values for 2020 to develop 2021-2030 growth surrogate data for this category.

Category 23. SCC 10100601; SIC code 4911 - Electric Generation, Natural Gas, Boilers > 100 MMBtu except Tangential

For this category, Pechan used the same growth surrogate data identified for category 22 above.

Category 24. CES 74682 - Cogeneration, Fuel Unspecified

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the total amount of fuel consumed in generating cogeneration energy.

To assist in developing growth surrogate data for this category, Pechan compiled 2000-2020 projections for total cogeneration (in billion kilowatt hours) for the Western Systems Coordinating Council/California-Southern Nevada (WSCC/CSN) from the AEO (DOE, 1998). The WSCC/CSN area approximates the State of California. Pechan developed a multiple regression analysis based on the AEO data and potential surrogate indicators to identify REMI variables that correlate with the AEO cogeneration data. The regression analysis indicated that Electric Utility output and Durables Manufacturing sector employment are statistically correlated with these cogeneration data (adjusted $r^2 = 0.993$; t-statistics of 54.18 and 4.132, respectively). To develop county-level growth surrogate data for this category, Pechan employed REMI data for 1970-2030 and the equation identified from the regression analysis.

Category 25. CES 82081 - I.C. Reciprocating Engines, Gaseous Fuel (Unspecified) (Oil and Gas Production)

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the total amount of natural gas consumed in reciprocating internal combustion engines that are used in the oil and gas production sector.

To develop 1970-1996 growth surrogate data for this category, Pechan used REMI county-level output data for SIC code 13 (Oil and Gas Extraction). For 1997-2020, Pechan used county-level REMI output data for SIC code 13 adjusted by National energy adjustment factors for the Mining sector. These mining sector energy adjustment factors were based on AEO forecasts. Pechan used the same approach to develop 2021-2030 growth surrogate data, with the exception that Pechan extrapolated the 2021-2030 energy adjustment factors based on the 2000-2020 Mining sector energy adjustment factors.

Category 26. SCC 20200202; SIC code 1311 - I.C. Engines, Reciprocating, Natural Gas (Industrial)

This is a point source category that is not included in ARB's Methods Manual. Based on the SCC and SIC code, the emissions activity for this category is the amount of natural gas burned in reciprocating internal combustion engines used within SIC code 1311.

To represent 1970-1996 growth surrogate data, Pechan used county-level REMI output data for SIC codes 131 and 132 (REMI does not report separate data for SIC code 131). For 1997-2020 growth surrogate data are based on REMI output data for SIC 131-2 after adjusting for the projected change in National natural gas use per output in the Mining sector as reported in the AEO (DOE, 1998). Pechan used an analogous approach to estimate 2021-2030 activity. For these years, however, Pechan employed National Mining sector natural gas intensity adjustments based on extrapolations of the adjustment factors developed from the 2000-2020 AEO projections.

Category 27. SCC 30600106; SIC code 2911 - Process Heaters, Process Gas-Fired (Petroleum Industry)

This is a point source category that is not included in ARB's Methods Manual. Based on the SCC and SIC code, the emissions activity for this category is the amount of process gas combusted in process heaters that are used in SIC code 2911.

For 1970-1996, Pechan used REMI SIC code 291 county-level output to represent growth surrogate data for this category. For 1997-2020, Pechan employed county-level REMI output data for SIC 291 after applying National adjustment factors representing the change in the amount of still gas consumed per unit of output in the refinery industry. These National adjustment factors were developed from data reported in the AEO (DOE, 1998). For 2021-2030, Pechan used the same approach as for 1997-2020 with the exception that Pechan developed National still gas energy adjustment factors based on extrapolations of the 2000-2020 AEO data.

Category 28. CES 66787 - Internal Combustion (I.C.) Reciprocating Engines, Natural Gas (Manufacturing and Industrial)

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of natural gas consumed in reciprocating internal combustion engines in the Manufacturing sector.

Growth surrogate data for 1970-1996 for this category are based on REMI output in the Manufacturing sector. To develop growth surrogate data for 1997 through 2030, Pechan used REMI output data for the Manufacturing sector after adjusting for the National change in natural gas use per unit of output in the Industrial sector. Data to develop the National change in natural gas use per unit of Industrial sector output were available from the AEO through 2020 (DOE, 1998). Pechan extrapolated the National change in Industrial sector natural gas intensity through 2030 based on the available AEO data.

**Category 29. CES 47142 - Fuel Combustion, Other, Natural Gas
(Manufacturing and Industrial)**

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the total amount of natural gas consumed in the Manufacturing sector.

Pechan used the same data as described above for CES 66787 (category 28) to represent the growth surrogate data for this category.

**Category 30. CES 83071 - Fuel Combustion, Other, Residual Oil
(Manufacturing and Industrial)**

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the total amount of residual oil consumed in the Manufacturing sector.

For 1970-1996 growth surrogate data, Pechan compiled REMI county-level output data for the Manufacturing sector. For 1997-2020, Pechan compiled REMI county-level output projections for the Manufacturing sector adjusted for the change in National residual oil consumption per dollar of output in the Industrial sector, which is available from the AEO (DOE, 1998). To develop emissions activity data for 2021-2030, Pechan computed composite REMI/AEO data. The energy intensity data for 2021-2030 was extrapolated from the 2000-2020 AEO energy intensity data.

**Category 31. CES 83998 - Agricultural Irrigation, I.C. Engines,
Diesel/Distillate Oil (Food and Agricultural Processing)**

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of distillate oil consumed by I.C. engines used in providing irrigation for agriculture.

Pechan used REMI Farm sector output as the growth surrogate for this category for 1970-1996. For 1997-2020, Pechan used county-level REMI output data for the Farm sector adjusted for the change in National distillate fuel intensity for the Agricultural sector as developed from AEO data (DOE, 1998). For 2021-2030, Pechan used the same approach as for 2000-2020, with the exception that the change in National distillate fuel intensity for the Agricultural sector was extrapolated from the 2000-2020 change in distillate fuel intensity reported in the AEO.

**Category 32. CES 47233 - Fuel Combustion, Other, Liquid Fuel
(Unspecified) (Food and Agricultural Processing)**

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the total amount of food and agricultural processing-based petroleum-related fuel combustion.

For 1970-1996, Pechan compiled REMI Farm sector output data as the growth surrogate for this source category. For 1997-2020, Pechan compiled county-level REMI output data for the Farm sector adjusted for the projected change in National petroleum energy intensity for the Agricultural sector. The National change in total petroleum intensity for the Agricultural sector was developed from projection data reported by the AEO (DOE, 1998). Pechan used an analogous approach to develop growth surrogate data for 2021-2030, with the exception that the National change in Agricultural sector petroleum intensity was extrapolated from the 2000-2020 change in petroleum intensity.

Category 33. CES 47167 - Fuel Combustion, Other, Natural Gas (Service and Commercial)

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of natural gas combustion occurring in the Service and Commercial sector.

To represent 1970-1996 growth surrogate data for this category, Pechan employed REMI county-level output data for the Services sector. For 1997-2030, Pechan utilized REMI Services sector output adjusted for the National change in natural gas intensity for the Commercial sector. The 1997-2020 change in natural gas intensity was developed from projected AEO data (DOE, 1998); the 2021-2030 natural gas intensity change was extrapolated from the 2000-2020 change derived from the AEO projections.

Category 34. CES 47159 - Fuel Combustion, Other, Distillate Oil (Service and Commercial)

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of distillate oil combusted in the Service and Commercial sector.

For 1970-1996 emissions activity for this source category, Pechan used REMI county-level output projections for the Services sector. For 1997-2030, Pechan used REMI county-level output projections for the Services sector adjusted for the National change in actual (1997-2020) and extrapolated (2021-2030) distillate fuel use per dollar of Commercial sector output. The National change in distillate fuel use per dollar of output was developed from AEO data (DOE, 1998).

Category 35. CES 57281 - Class II and III Landfills, Municipal Solid Waste

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of municipal solid waste in place in Class II and III landfills. The definitions for Class II and Class III landfills are based on landfill design (e.g., single liner) and siting criteria.

IWMB has county-level total landfill tonnage estimates for 1990-1999 (these estimates are not provided by landfill class). In addition, the IWMB has state-level estimates for 1989 (pre-

1989 data are not available because the impetus for collecting these data was the Integrated Waste Management Act of 1989). Pechan used the county-level tonnage estimates as the 1990-1999 growth surrogate data for this category.

Pechan conducted an analysis of the IWMB statewide landfill tonnage data regressed against the following REMI variables: population; employment and output in the Water and Sanitation Services sector; and total gross regional product (i.e., the total value of the State's economy). The regression analysis identified a strong statistical correlation between landfill tonnage and both population and total gross regional product (adjusted $r^2 = 0.964$; t-statistics of -4.095 and 11.54, respectively). Based on this analysis, Pechan used the established equation and the historical and forecast REMI data for these two variables to develop 1970-1990 and 1999-2030 landfill tonnage estimates. Pechan then developed county-level ratios of the 1990 tonnage estimates to the pre-1990 tonnage estimates from this equation. Pechan applied these ratios to the actual 1990 tonnage values to develop pre-1989 growth surrogate data. Pechan implemented an analogous procedure to develop post-1999 growth surrogate data.

Category 36. CES 83659 - Degreasing, Cold Cleaning (Batch, Conveyor, Spray Gun), Petroleum Naphtha

This source category is not discussed in ARB's Methods Manual; Pechan assumes that the amount of petroleum naphtha used in cold cleaning operations is the emissions activity for this source category.

No long-term California emissions activity data were identified for this category. To identify surrogate growth indicators for this category, Pechan compiled National special naphtha total demand and National special naphthas demand for cold cleaning data for 1987, 1989, 1992, 1993, 1996, 1998, 2003, and 2008 from two Freedonia Group, Inc. (Freedonia) reports – "Solvents to 2001," and "Solvents to 2003" (Freedonia, 1997 and Freedonia, 1999b). Pechan then conducted a regression analysis that included Freedonia's special naphthas demand for cold cleaning data as the dependent variable. Pechan regressed the Freedonia data against employment and output data for the following surrogate indicators, which were identified as major degreasing/cold cleaning industries in a solvent cleaning/degreasing report produced for ARB by Pechan and/or the "Solvent Cleaning" chapter of the EIIP emission estimation report: Automobile Parking, Repair, and Services; Durables Manufacturing; and Total Manufacturing (Roe, 1996 and ERG, 1997a). Pechan used the National data for these sectors from the REMI standard model in the analysis. Based on the best-fit equation identified in the multiple regression analysis (adjusted $r^2 = 0.910$; t-statistics for Automobile Parking, Repair and Services of -3.488 and of Total Manufacturing output of 2.310), employment data for the Automobile Parking, Repair, and Services sector and Total Manufacturing output data are correlated to special naphthas demand for cold cleaning.

Pechan compiled 1970-2030 output from the identified equation using county-level REMI inputs for Automobile Parking, Repair and Services employment and Total Manufacturing output data. A review of the equation output identified very large values in the early years of the study period. The cause of this phenomenon is the major decline in special naphtha demand based on the Freedonia data that were used in the regression analysis. Because the regression approach

resulted in unreasonable values for certain years, Pechan applied a no growth assumption in developing growth surrogate data for years beyond the period included in the regression analyses (1987-2008). The no growth assumption was utilized because the most strongly correlated variable was identified as negatively correlated with naphtha demand.

Category 37. CES 58685 - Natural Gas Transmission Losses

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of natural gas transmitted.

Pechan was unable to locate information on the amount of California natural gas transmission activity. Pechan used the natural gas production data developed for CES 46441 to represent the growth surrogate data for this category (see the discussion for Category 04 above for details on how these data were developed).

Category 38. SCC 30500606 and SIC code 3241 - Cement Manufacturing (Dry Process)

This source category is not discussed in ARB's Methods Manual; Pechan assumes that the emissions activity for this category is based on the amount of dry process cement produced within SIC code 3241.

Pechan obtained portland cement production data for the State of California for the period 1980-1997 from the U.S. Geological Survey's web site (for 1992-1997) and the Bureau of Mines' "Minerals Yearbook, Mineral Industry Surveys; Volume II: Domestic" for earlier years (USGS, 2000 and BOM, 1992). Although data on masonry cement production were withheld from publication for most years due to confidentiality concerns, this omission has little impact on the total cement production estimate. For one year where masonry cement production was available (1996), masonry cement accounted for less than 2 percent of total cement production. Although the cement production data do not distinguish between dry and wet process, all production is assumed to be dry process because the U.S. Geological Survey indicates that all clinker plants in California used the dry process in 1994.

Pechan regressed 1980-1997 California portland cement production against California data from the REMI models for the following variables: employment and output in the Construction sector and employment and output in the Hydraulic Cement Manufacturing sector. A statistically significant relationship was identified between portland cement production and output from the Hydraulic Cement Manufacturing sector (adjusted $r^2 = 0.739$; t-statistic = 7.0). To develop growth surrogate data for this category, Pechan utilized the output of the identified equation relating portland cement production to output from the Hydraulic Cement Manufacturing sector.

Category 39. CES 85654 - Architectural Coatings, Oil-Based Industrial Maintenance Coating

This source category is not discussed in ARB's Methods Manual; Pechan assumes that the emissions activity for this category is the amount of solvents used in oil-based industrial maintenance coatings.

Because no California-specific information was identified for this category, Pechan compiled National production data for industrial maintenance paints from the Bureau of the Census publication *Census of Manufacturers*, "Paint and Allied Products" (Census, 1997b). Production data for these specific paints were only available for 1982-1984. However, production data were available from the Bureau of the Census for 1984-1998 for "industrial new construction and maintenance paints." Pechan applied the rate-of-change in the Census' 1984-1998 series to the Census' 1984 industrial maintenance paint value to estimate the post-1984 trend in industrial maintenance paints production over this period.

To identify surrogates relating to industrial maintenance paints production, Pechan regressed the National production series against National output and employment data for the following variables: Total Manufacturing; Durables Manufacturing; Chemicals Manufacturing; and Petroleum Refining. These last two variables were included because *The Rauch Guide to the U.S. Paint Industry, 1998-1999 edition*, indicated that these industries account for approximately one-half of the consumption of industrial maintenance paints (IMC, 1997). The identified best-fit equation includes output for both the Chemicals and Petroleum Refining sectors (adjusted $r^2=0.860$; t-statistics are 5.381 and -2.262 for Chemicals output and Petroleum Refining output, respectively). Pechan used REMI output data for these sectors and the identified equation to develop estimated growth surrogate data for this category. Based on the unreasonably large 1970-2030 growth rates from the resulting estimates and the limitations of the data used in the regression analysis (no California-specific data and lack of long-term data specific to maintenance paints), Pechan decided not to use the regression output in developing the growth surrogate data for this category. Instead, Pechan compiled REMI output data for the Chemicals sector. These data represent the most strongly correlated variable based on the regression analysis.

The 1998-1999 *Rauch Guide* states that solvent-based coatings account for nearly 75 percent of the total volume of industrial maintenance coatings. The *Rauch Guide* also notes that water-based paints "are growing at an above average rate, although they are not as efficient as solvent-based materials." The *Rauch Guide* notes that improved water-based formulations are continuing to be developed, so growth should be above average. However, Pechan was unable to identify any long-term data or forecasts on the amount of solvents used in industrial maintenance paints. Based on reports that water-borne coatings are expected to make significant in-roads in industrial maintenance markets and an estimate that the overall U.S. market for solvent-borne coatings is growing at one-half the overall coatings industry growth rate, Pechan assumed that post-1998 growth in oil-based industrial maintenance paints will occur at one-half the rate of total industrial maintenance paints (Hoffman, 1997 and Boswell, 1998). Pechan analogously adjusted the REMI Chemicals sector data for any areas with projected increases in Chemicals output. All pre-1998 emissions activity data were derived directly from the REMI output data in the Chemicals sector.

Category 40. CES 85852 - Architectural Coatings, Water-Based Flat Coatings

This source category is not specifically discussed in ARB's Methods Manual; Pechan assumes that the amount of water-based flat architectural coatings consumed is the emissions activity for this category.

Pechan was unable to identify long-term water-based flat architectural coatings consumption data. Pechan was able to compile 1985-1998 National data on production of water-based interior flat architectural coatings and water-based total exterior architectural coatings from the Bureau of the Census (Census, 1997a). To identify a growth surrogate for this source category, Pechan regressed the sum of these two data series for each year over the 1985-1998 period against National data for the following variables: output and employment in the Construction sector; housing expenditures; value of new commercial construction; population; and output and employment in the Real Estate industry. The multiple regression analysis indicated the best statistical fit with employment in the Real Estate industry (adjusted $r^2 = 0.798$; t-statistic = 7.23). Based on this analysis, Pechan used the Real Estate industry employment data from REMI's economic model in the identified equation to develop county-level growth surrogate data for this source category.

Category 41. CES 46870 - Cutback Asphalt

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of cutback asphalt used in road paving.

ARB contacted Caltrans to inquire about the availability of historical and projected data on cutback asphalt use. Caltrans notified Pechan that such data are not readily available. In keeping with previous ARB practice, Pechan used county-level employment in the Construction industry (SIC codes 15-17) to represent growth in this category.

Category 42. CES 66605 - Livestock Wastes

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the number of head of livestock in California.

Pechan compiled county-level 1970-1998 data on the number of head of cattle and calves, sheep and lambs, and hogs from the CDFA (CDFA, 1988 and CDFA, 1998). (It should be noted that Pechan also investigated the availability of data on the number of chickens from the CDFA, but found that the available data were not comprehensive nor available for a similar time frame as were the other livestock data.) Pechan prepared multiple regression analyses of these data with the following potential explanatory variables: employment and output in the Farm sector; employment and output in the Meat Products sector; employment and output in the Dairy Products sector. The analysis indicated that none of these regressors provide a statistically significant correlation with the number of livestock. This phenomenon is a result of the fact that

the State total number of head of these livestock has been fairly consistent throughout the period. Therefore, Pechan employed a no growth assumption to represent trends in this source category.

Category 43. CES 47241 - Agricultural Burning-Prunings

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of agricultural prunings burned represents the emissions activity for this source category.

The "Open Burning" section of EPA's *AP-42* document lists emission factors for the following orchard crops: almond, apple, apricot, avocado, cherry, citrus (orange, lemon), date palm, fig, nectarine, olive, peach, pear, prune, walnut, and unspecified (EPA, 1999). Based on this information, Pechan compiled historical acreage data for the following CDFA commodity codes: 261999 (Almonds, All), 211999 (Apples, All), 217999 (Apricots, All), 221999 (Avocados, All), 213199 (Cherries, Sweet), 208059, 201119 (Oranges, Navel), 201519 (Oranges, Valencias), 201999 (Oranges, Unspecified), 204999 (Lemons, All), 224999 (Dates), 225999 (Figs, Dried), 218199 (Nectarines), 226999 (Olives), 212199 (Peaches, Freestone), 212399 (Peaches, Clingstone), 212999 (Peaches, Unspecified), 214199 (Pears, Bartlett), 214899 (Pears, Asian), 214999 (Pears, Unspecified), 215999 (Prunes, Dried), 263999 (Walnuts, English), and 265999 (Walnuts, Black). Pechan compiled these data from data provided by the CDFA (CDFA, 2000). After compiling harvested acreage data for these commodity codes, Pechan identified seven counties in ARB's 1996 inventory with emissions for this category for which no CDFA acreage data are reported for the above crops.

Based on this information, the ARB Contract Manager agreed that Pechan would use total harvested non-pasture crop acreage data to represent the growth surrogate data for this category rather than the acreage data for crops with related emission factors in EPA's *AP-42* document. Therefore, Pechan used the data developed for category number 19 above (CES 47431) to represent the growth surrogate data for this category.

Category 44. CES 47258 - Agricultural Burning-Field Crops

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of agricultural field crops burned.

EPA's *AP-42* document lists the following field crops with emission factors for agricultural burning: asparagus, alfalfa, barley, beans (red), corn, cotton, hay (wild), oats, pea, pineapple, rice straw, safflower, sorghum, sugar cane, wheat, and field crops, unspecified (EPA, 1999).

Pechan compiled county-level 1986-1998 harvested acreage data from the CDFA for the following crop categories: 302199 (Asparagus, Fresh Mkt), 302299 (Asparagus, Proc), 302999 (Asparagus, Unspecified), 181999 (Hay, Alfalfa), 113944 (Barley Malting), 113995 (Barley, Feed), 113999 (Barley, Unspecified), 161717 (Beans, Red Kidney), 161722 (Beans, Sm Red), 111559 (Corn, White), 111991 (Corn for Grain), 111992 (Corn for Silage), 111998 (Corn, Crazy), 121219 (Cotton Lint, Upland), 121229 (Cotton Lint, Pima), 121299 (Cotton Lint,

Unspec.), 188799 (Hay, Wild), 112999 (Oats for Grain), 361199 (Peas, Green Fr. Mkt.), 361299 (Peas, Green, Processing), 361999 (Peas, Green, Unspecified), 362999 (Peas, Cowpea & Blackeye), 394999 (Peas, Edible Pod), 198199 (Rice, Wild), 158269 (Safflower), 114991 (Sorghum, Grain), 114992 (Sorghum, Silage), 131999 (Sugarcane), 101999 (Wheat All), and 198999 (Field Crops, Unspec.). These data were compiled from data provided by the CDFA (CDFA, 2000). The total State acreage for these crops was regressed against 12 potential explanatory variables. The best equation was identified with population and employment in the Farm and Garden Machinery Equipment sector (adjusted $r^2 = 0.883$; t-statistics of -8.506 and 3.538, respectively). The negative coefficient for population can likely be explained by increased urbanization resulting in less acreage utilized for agricultural use.

Pechan used REMI population and Farm and Garden Machinery Equipment employment data in the equation described above to develop state-level factors to apply to the actual 1986 and 1998 acreage values. A review of the resulting estimates indicated near zero values by the year 2030. Because it is unclear that this assumption is reasonable, Pechan implemented a no growth assumption to represent the growth surrogate data for this category for all non-regression analysis years (i.e., all pre-1986 and post-1998 years); the regression output was used for 1986-1998.

Category 45. CES 47282 - Range Improvement

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is the amount of rangeland burned to enhance the growth of grass for cattle.

Because Pechan was unable to identify long-term historical data on the amount of rangeland burned in California, the ARB Contract Manager concurred with Pechan's recommendation to use the number of acres of rangeland to represent the historical emissions activity data for this category. Pechan compiled 1988-1998 county-level pasture acreage data from CDFA. Pechan regressed these data at the state-level with the following variables: employment and output in the Farm sector; employment and output in the Meat Products sector; employment and output in the Food Manufacturing sector; and population. The best statistical fit was identified with two variables: population and employment in the Food Manufacturing sector (adjusted $r^2 = 0.880$; t-statistics of -8.555 and 2.714, respectively). As noted for other agriculture-related categories, the negative coefficient for the population variable can likely be explained by increased urbanization resulting in less acreage being used for rangeland.

Pechan used the REMI population and Food Manufacturing sector employment data in the equation described above to develop growth surrogate data for this category.

Category 46. CES 47274 - Forest Management

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the number of acres of prescribed burning activity for forest management is the emissions activity for this category.

Pechan obtained data from the California Department of Forestry and Fire Protection (CDFFP) for 1982-1999 on the number of Federal and private land acres burned by each of 27 regions (40 counties are included in these regions) in the State. These data exclude fire acreage data for numerous non-CDFFP jurisdictions. Specifically, the CDFFP data do not account for fires occurring on U.S. Forest Service, State Parks, National Parks, and Indian lands. Because of the level of effort associated with contacting several entities to obtain these data, along with the effort required to compile and reconcile the available data from each source, the ARB Contract Manager agreed with Pechan's assessment that it was infeasible to develop historical fire acreage data for this category. Pechan implemented a "no growth" assumption for this category because the level of emissions activity is not expected to trend with any socioeconomic variable.

Category 47. CES 47266 - Weed Abatement

Individual Air Districts have the responsibility for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of agricultural weeds burned. No data were located characterizing the level of this emissions activity. Pechan assumes that total non-pasture crops harvested acreage data excluding acreage from orchards and vineyards can be used to represent the growth surrogate data for this category. See the category 20 (CES 833337) discussion for a description of how these data were compiled from regression output.

Category 48. CES 47290 - Non-agricultural Open Burning

Individual Air Districts are responsible for estimating emissions for this category. Pechan assumes that the amount of non-agricultural debris burned is the emissions activity for this source category.

Pechan was unable to identify data characterizing the historical emissions activity for this category (e.g., Pechan contacted ARB to identify whether permit/compliance data were available that would represent activity levels for this category). Pechan assumed no emissions activity growth for this category for all counties where 80 percent or greater of the 1990 county's population was classified in urban areas based on Bureau of the Census estimates. This approach results in 25 counties with a no growth assumption applied. For the remaining counties, Pechan used REMI county-level population data as the surrogate growth indicator. Note that the 80 percent threshold was selected because this is EPA's current assumption for use in estimating open burning emissions (i.e., counties in the Nation with at least 80 percent of their population in urban areas are assumed to have no open burning emissions).

Category 49. CES 60418 - Commercial Charbroiling

Individual Air Districts are responsible for estimating emissions for this category. Pechan assumes that the emissions activity for this category is based on the amount of meat that is commercially charbroiled.

Pechan identified that the San Diego Air District estimated the emissions activity for this category based on the total number of food service facilities (from food service inspection records) and the assumption that 1 out of every 10 such facilities does charbroiling. Because the *Census of Retail Trade* is published every 5 years, Pechan compiled data on the number of “refreshment places” as reported in this publication for each 5 years in the period 1972-1997 (Census, 1997d). Refreshment places were identified as the activity data that was expected to most closely approximate the trend in the number of establishments where charbroiling takes place. As defined by the Census, this category includes “establishments that prepare items such as chicken and hamburgers for consumption either on or near the premises or for “take-home” consumption. (Because the 1997 *Census of Retail Trade* no longer uses the same categorization for the eating and drinking sector as previous sectors, Pechan used the number of establishments in the “limited-service restaurants” category, which includes “establishments primarily engaged in providing food services where patrons generally order or select items and pay before eating.”)

Next, Pechan regressed the 1972-1997 number of refreshment places in California against REMI California output and employment data for the Eating and Drinking Places sector (SIC code 581). The regression analysis indicated that output from this sector provided a strong correlation with the number of refreshment places (adjusted $r^2=0.957$ and t-statistic = 10.55). Pechan used the identified equation and REMI output data for the Eating and Drinking Places sector to develop growth surrogate estimates for this category.

Because the regression-based estimates resulted in extremely large growth rates over the study period, Pechan developed the final growth surrogate data by combining the regression output for 1972-1997 (the years for which the regression analysis was conducted) and the county-level trend in Eating and Drinking Places sector output. This sector’s output was used to reflect pre-1972 and post-1997 trends for this category because it was identified by the regression analysis as statistically correlated with the number of refreshment places.

Category 50. SCC 20200202; SIC code 4922 - I.C. Engines, Industrial, Natural Gas, Reciprocating

As a point source category, this category is not discussed in ARB’s Methods Manual; Pechan assumes that the emissions activity for this category is based on the amount of natural gas combusted in internal combustion engines used in SIC code 4922 establishments.

To assist in developing growth surrogate data for this category, Pechan conducted a multiple regression analysis of state-level CEC natural gas consumption data for the Transportation sector within SIC codes 491-493, 496, 498. Based on this analysis, Pechan identified a statistically significant relationship between the CEC natural gas consumption data and the following two explanatory variables: employment in the Gas Utilities sector (SIC code 493 and part of SIC code 492) and output in the Crude Petroleum, Natural Gas, and Gas Liquids sector (adjusted $r^2 = .907$; t-statistics of 3.183 and 5.583, respectively). To develop growth surrogate data for this category, Pechan used the equation determined from the regression analysis and REMI data for the two explanatory variables.

D. ELECTRIC UTILITY-RELATED CATEGORIES

After initiating the analysis of growth surrogates for the 50 ARB-identified categories, ARB requested that Pechan incorporate CEC data as a surrogate growth indicator for all electric utility-related point sources. This section describes the final approach used to develop growth surrogate data for these sources.

For any point source category that is assigned an SIC code of 4911 or 4931, Pechan selected a CEC fuel use data series as the 2000-2020 growth surrogate data. As identified in the below table, Pechan selected the appropriate data series based on each record's EIC and SCC:

EIC	SCC	CEC Electric Generating Fuel Use Data Series
<=99999999999999	Any natural-gas using	Natural gas fuel use
<=99999999999999	All excluding natural gas	Total fuel use excluding natural gas
> 99999999999999	Any	Total fuel use

For the first row above, Pechan used the approach described for category number 22 (SCC 10100601; SIC code 4931) in Section III.C.4 to develop pre-2000 and post-2020 data for natural gas combustion point sources with SIC codes 4911 and 4931.

For the inventory records associated with the second row in the above table, Pechan first compiled CEC data for 2000-2020 reflecting total non-natural gas fuel use data by Air Basin/District/county. Pechan then compiled data for all Air Basin/District/county combinations for which CEC did not report data, but for which Air Basin-level data were available. Pechan assigned Air Basin-level data to these Air Basin/District/county combinations. Pechan applied a no growth assumption to develop growth surrogate data for 2021-2030 for this category. This assumption was employed because there is no change in the CEC fuel use values for this category over the 2010-2020 period. For pre-2000 estimates, Pechan applied state-level growth rates to the 2000 CEC fuel use data. Pechan used DOE electric utility non-natural gas energy consumption data to represent 1970-1975 trends. For 1976-1998, Pechan used CEC's state-level electric generation fuel consumption data excluding natural gas. For 1999, Pechan assumed no growth based on the less than one percent average annual growth rate in non-natural gas electric utility fuel use over the 1976-1999 period.

For the third row above, Pechan first compiled CEC data for 2000-2020 reflecting total electric utility fuel use data by Air Basin/District/county. Pechan followed a similar approach to that described above for the other electric utility-related growth parameters in estimating data for missing Air Basin/District/county combinations. For 2021-2030, Pechan used the change in REMI Electric Utility sector output adjusted by an equation relating CEC 2000-2020 electric generation total fuel use data to Electric Utility sector output. This equation was developed from a multiple regression analysis that indicated a strong statistical relationship between the two variables (adjusted $r^2 = 0.985$, t-statistic = 35.90). For pre-2000 growth surrogate data, Pechan applied state-level growth rates to the 2000 CEC total fuel use data. Pechan utilized DOE electric utility energy consumption data to represent growth rates over the 1970-1975 period (DOE, 2000). For 1976-1998, Pechan employed CEC total electric generation fuel consumption

data (Ewing, 2000). For 1999, Pechan used the average annual growth rate over the 1976-1998 period (approximately 1.5 percent) calculated from the historical CEC data.